

# **MANUEL CANADIEN D'ÉVALUATION DES INCIDENCES SUR LA SANTÉ**

---

## **Volume 3**

### **Rôle du professionnel de la santé**

**Version de travail/Ne pas citer  
Décembre 1999**

Ce document a été divisé dans une série de fichiers pour faciliter leur téléchargement de notre site du web.

Partie 7 de 9

## **BIostatistique : DÉFINITIONS ET MÉTHODES**

### **Table des matières**

**Introduction**

**Biostatistique et études d'impact sur la santé**

**Analyse statistique**

**Description des données**

**Estimation**

**Test d'hypothèse**

**Observation des effets statistiquement significants**

**Comparaison de groupes**

**Analyse de régression**

**Analyse multivariée**

**Sujets particuliers**

**Estimation de l'activité**

**Toxiques avec dose-seuil**

**Toxiques sans dose-seuil**

**Extrapolation de l'animal à l'homme**

**Contraintes liées à l'utilisation de la biostatistique**

**Conclusions**

**Ressources documentaires, spécialistes, organisations**

**Bibliographie**

## Introduction

On utilise la biostatistique dans les études d'impact sur la santé et les évaluations environnementales afin de décrire quantitativement l'état de santé d'une population ou sa condition économique et sociale, ainsi que les effets prévus du projet à l'étude. Mais que sont les statistiques et dans quelles circonstances les utilise-t-on?

Une statistique est la mesure globale d'un aspect quantifiable des données. Souvent, les données observées constituent un sous-ensemble ou un *échantillon* d'une *population* de données. Il est plus simple, et moins coûteux, d'étudier un échantillon qu'une population. Or, l'objectif d'une étude statistique est généralement de faire des inférences sur la population. Il est possible de réaliser cet objectif au moyen d'un échantillon, pourvu que celui-ci soit *représentatif* de la population étudiée. L'essentiel est de comprendre la nature du rapport entre l'échantillon et la population. Les façons de prélever une échantillon sont nombreuses, voire infinies. Ainsi, tout énoncé quantifié portant sur l'échantillon devient un jugement *probabiliste* sur la population. Le prélèvement de l'échantillon, l'analyse des données et la formulation d'inférences sur la population sont les trois étapes d'une opération complexe qui fait l'objet d'une discipline appelée *statistique*.

Les statistiques ont pour principaux avantages d'être représentatives de la population et d'être reproductibles jusqu'à un certain point. On garantit la représentativité des données en prélevant aléatoirement un échantillon dans la population. Par exemple, lorsqu'on décontamine un site de déchets

Une *statistique* est la mesure globale d'un caractère quantitatif des données. Par exemple, la moyenne est la *valeur moyenne* des données.

### Définition :

La *statistique* est une discipline qui comprend un ensemble de techniques et de règles devant conduire à l'acquisition, à l'analyse et à la mise en forme de données, et à la formulation d'inférences à partir d'un échantillon au moyen de la théorie des probabilités. Régie par les *lois de probabilité*, cette discipline implique des données numériques et la dimension aléatoire, mais exige par ailleurs

On appelle *population* l'ensemble des objets auxquels s'intéresse une étude.  
On appelle *échantillon* une partie de cette population.

dangereux, il est utile de connaître la concentration de contaminants sur le site. Or, comme il n'est pas possible de mesurer les concentrations sur tout le site, on procède à un échantillonnage. Si cette opération est bien faite, on obtient généralement des données qui reflètent fidèlement le niveau de contamination du site. La randomisation dans l'échantillonnage, qu'elle soit soumise ou non à des contraintes, est indispensable pour « *Risque* : danger pour la santé résultant de l'exposition à une substance, un processus ou un produit, et probabilité que ce danger se matérialise. » (Santé Canada, 1997) garantir la représentativité. On peut même se prononcer sur la probabilité que la fonction des observations de l'échantillon corresponde à la fonction des observations théorique de la population en se servant des intervalles de confiance.

Les mesures relatives à la population d'individus étudiée varient, c'est-à-dire qu'elles ne sont pas des valeurs fixes prévisibles. Il existe donc une *distribution* statistique pour ces mesures. L'*histogramme* d'un échantillon de mesures reproduit assez fidèlement la distribution d'un caractère dans la population. L'une des distributions les plus courantes est la distribution normale, représentée par une courbe en cloche symétrique.

On décrit les distributions par des paramètres statistiques fondamentaux. Par exemple, la distribution normale se définit entièrement par la moyenne et la variance; en d'autres termes, on estime la distribution de population en calculant l'estimation de la moyenne et de la variance de l'échantillon de mesures. En définissant les distributions de population par un type de distribution et les paramètres statistiques qui le caractérisent, nous sommes en mesure de comparer des populations entre elles.

Il n'est pas toujours simple de former un échantillon représentatif. Le mode de collecte des

La ran Chapitre D  
Table des matières  
do Chapitre D  
Table des matières  
misation est une  
c o n d i t i o n  
essentielle pour  
p r é l e v e r u n  
échantillon qui soit  
représentatif de la  
p o p u l a t i o n .

données peut être assujetti à des contraintes. On élabore donc des plans de sondage spéciaux afin de formuler le plus efficacement les inférences les plus justes. Dans le domaine des plans d'expérience, le plan randomisé à blocs et le plan randomisé à blocs incomplets sont des exemples de protocoles normalisés dont le choix dépend des objectifs et des contraintes.

L'application de méthodes statistiques nécessite habituellement la formulation d'hypothèses concernant les données. Par exemple, l'utilisation de la statistique t pour tester l'hypothèse que la moyenne d'une population est nulle sera valide uniquement si les données proviennent de la même distribution normale et si elles sont indépendantes entre elles. Si on applique des méthodes statistiques sans tenir compte des hypothèses sur lesquelles elles reposent, on risque de tirer des conclusions erronées et trompeuses.

Plus le protocole d'échantillonnage et les hypothèses à tester sont complexes, plus les méthodes d'analyse des données seront complexes. Les statisticiens de profession reçoivent une formation universitaire pour être capables d'appliquer convenablement des méthodes statistiques complexes, ce qui démontre que la statistique est un domaine d'études scientifiques important. Les statisticiens qui sont titulaires d'une maîtrise ou d'un doctorat ont une formation qui leur permet d'élaborer de nouvelles méthodes ou d'en adapter des anciennes quand il le faut.

Les personnes qui travaillent à des études d'impact sur la santé mais qui ne sont pas des statisticiens gagneraient à recevoir une formation en statistique pour comprendre les concepts, bien que le chercheur ou l'évaluateur ne soit pas nécessairement tenu de savoir comment appliquer les méthodes en tant que telles. Dans certains cas, il suffira de savoir à quel moment on doit consulter un statisticien. En ce qui concerne la formation proprement dite, elle peut prendre toutes sortes de formes : depuis le cours d'introduction

jusqu'au cours s'inscrivant dans un programme d'études supérieures. Les programmes d'études de nombreuses disciplines, telles que la toxicologie, la psychologie et l'économie, comprennent au moins un cours de statistique obligatoire.

L'analyse statistique est beaucoup plus accessible aux non-spécialistes depuis la diffusion de l'ordinateur. On trouve maintenant sur les ordinateurs personnels des programmes faciles à apprendre et à utiliser et peu coûteux, qui peuvent exécuter des analyses assez complexes. L'éventail va des programmes où il suffit de pointer-cliquer aux langages de programmation très perfectionnés, comme le Système d'analyse statistique (*Statistical Analysis System*) (SAS, 1990) et S-Plus (MathSoft, 1990). Rappelons toutefois qu'il faut utiliser ces programmes avec circonspection afin de garantir l'intégrité des résultats de l'analyse.

### **Biostatistique et études d'impact sur la santé**

Les usages de la statistique sont nombreux dans les études d'impact sur la santé. Notons les plus courants : évaluation de l'exposition; évaluation du comportement des contaminants; évaluation du risque; coûts prévus pour atténuer l'impact; évaluation de l'exposition de base par rapport à l'exposition prévue; estimation de l'ampleur de l'impact; estimation de la population touchée; quantification des indicateurs de la santé; et teneur en produits chimiques toxiques dans les tissus humains (voir vol. 1).

L'utilisation de statistiques fiables dans les études d'impact sur la santé comporte plusieurs avantages :

- L'utilisation des méthodes statistiques appropriées hausse la crédibilité des résultats;

- On peut extrapoler sur des données d'échantillon pour faire des inférences sur une population et on peut définir à cette fin un niveau de confiance;
- La statistique permet d'évaluer la variabilité des résultats et met en évidence les inexactitudes dans les données.
- On peut estimer les "moyennes" étant donné la forme et la position de la distribution d'échantillonnage;
- On peut interpréter les résultats compte tenu de la variabilité estimée.

L'utilisation des statistiques a aussi ses limites :

- Même l'usage le plus minutieux des statistiques ne garantit pas que l'on obtiendra toutes les réponses voulues. Les statistiques ont un caractère foncièrement probabiliste et elles comportent donc un risque d'erreur (erreur de première espèce et erreur de seconde espèce).
- Les statistiques ne peuvent surpasser en qualité les données à partir desquelles elles sont établies. Voici les sources d'erreur possibles : la collecte des données ne se fait pas en conformité avec les objectifs particuliers de l'étude; les techniques de mesure ne sont pas fiables ou elles sont très variables; d'autres facteurs confusionnels peuvent aboutir à une mauvaise interprétation des résultats.
- L'analyse statistique dépend largement des hypothèses formulées. Si une ou plusieurs hypothèses s'avéraient invalides, les conclusions de l'analyse pourraient être faussées.

### Analyse statistique

On commet une **erreur de première espèce** lorsqu'on déclare qu'il existe une différence réelle entre deux échantillons, alors que la différence observée est due uniquement au hasard. (Traduction)

[Kennedy et Neville, 1974.]

On commet une **erreur de seconde espèce** lorsqu'on conclut qu'il n'existe pas de différence réelle entre deux échantillons, alors que cette différence existe vraiment. (Traduction)

[Kennedy et Neville, 1974.]

## Description des données

Avant d'analyser les données, il convient d'en déterminer la structure et le mode de collecte. Deux questions s'imposent :

- (1) Quel est le plan d'expérience ou le plan de sondage?
- (2) Quel genre de données ont été recueillies?

*Quel est le plan d'expérience ou le plan de sondage?*

Un plan élaboré adroitement comporte de nombreux avantages. Il peut faciliter l'application de méthodes statistiques justes et efficaces. Les résultats sont souvent plus faciles à analyser et à interpréter et ils sont plus facilement comparables avec ceux d'autres études.

Le plan d'expérience repose sur les principes statistiques suivants : randomisation, efficacité, bloc et structure. La randomisation garantit l'indépendance et la validité statistiques. Avec la randomisation, rien d'autre que l'expérience proprement dite ne doit caractériser les groupes expérimentaux. Par exemple, pour effectuer une randomisation conforme dans des expériences de toxicologie sur les animaux, on doit assigner aléatoirement les traitements (ou expériences) aux animaux et assigner aléatoirement ceux-ci à des cages. Par bloc on entend les blocs d'unités expérimentales. Par exemple, si des expériences étaient menées une journée quelconque et d'autres expériences, un autre jour, alors on désignerait le jour comme un effet de bloc. La structure désigne le schéma de l'expérience, par exemple nombre de facteurs, séquence temporelle pour les unités expérimentales.

*Quel genre de données ont été recueillies?*

De nombreuses observations peuvent avoir été recueillies durant l'expérience. On note trois types de résultats :

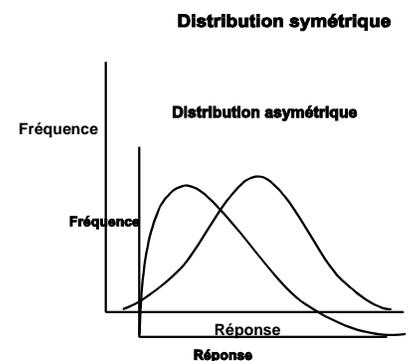
- données continues
- proportions
- données de dénombrement.

Le poids corporel (ou poids vif), la consommation d'aliments concentrés, les résultats d'expérience de chimie clinique ou d'hématologie et les résultats d'analyse d'urine, et le poids d'organe sont des exemples d'observations continues. Lorsqu'il s'agit de déterminer le nombre d'individus porteurs d'une tumeur parmi un groupe d'animaux, on parlera plutôt de proportion. Enfin, le nombre de globules blancs et le nombre de mutations dans une bio-analyse sont des exemples de données de dénombrement.

Il serait utile de pouvoir compter sur une méthode d'analyse statistique uniformisée, mais ce n'est malheureusement pas le cas. Certaines méthodes peuvent sembler convenir à de nombreuses situations, mais elles ne représenteront pas nécessairement le test le plus efficace ni n'auront la puissance voulue pour détecter les effets. C'est pourquoi il faut choisir le type d'analyse avec beaucoup de soin. Ce choix sera déterminé par le plan d'expérience et le type de données.

## Estimation

Une statistique est une fonction des données observées qui a pour objet de résumer un aspect particulier de la distribution des données, laquelle peut être symétrique, asymétrique, bimodale ou d'une autre forme. Dans le cas de la distribution normale, si l'on connaît la moyenne et la variance, on peut reconstituer entièrement la distribution. Il se peut que l'on ne connaisse pas

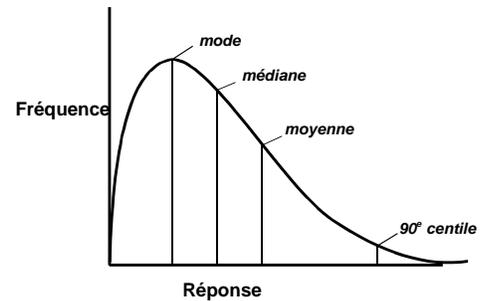


la moyenne et la variance vraies, mais on peut les estimer à l'aide des statistiques, c'est-à-dire en se servant de la moyenne et de la variance de l'échantillon, calculées à l'aide des données. La moyenne est une mesure de tendance centrale pour la distribution normale. Les autres mesures de tendance centrale sont la médiane, le mode et la moyenne géométrique. Les autres statistiques qui servent à décrire une distribution sont le percentile (ou quantile) et diverses mesures de variabilité telles que la variance et l'étendue des données.

<b>Définitions des statistiques les plus courantes</b>	
<b>Statistique</b>	<b>Définition</b>
<b>Moyenne</b>	<b>somme des observations divisée par le nombre d'observations</b>
<b>Médiane</b>	<b>valeur qui se trouve au centre d'un ensemble d'observations lorsque celles-ci sont rangées par ordre croissant ou décroissant</b>
<b>Moyenne géométrique</b>	<b><math>n</math>-ème racine du produit de <math>n</math> nombres non négatifs</b>
<b>Variance</b>	<b>somme des carrés des écarts de chaque observation par rapport à la moyenne, divisée par le nombre d'observations moins un</b>
<b>Écart type</b>	<b>racine carrée de la variance</b>
<b>Erreur-type de la moyenne</b>	<b>racine carrée de la variance divisée par le nombre d'observations</b>
<b><math>n</math>-ème centile</b>	<b>valeur dont <math>n</math> % des observations lui sont inférieures</b>
<b>Intervalle de confiance pour la moyenne à <math>n</math> %</b>	<b><math>n</math> % des intervalles obtenus contiendront la vraie valeur de la moyenne</b>

Il arrive souvent dans l'évaluation de l'exposition que l'on doive définir une mesure globale significative de l'exposition si les données proviennent d'une distribution asymétrique. En effet, si ces données contiennent un petit nombre de valeurs très élevées, ce qui est fréquent dans le cas des données sur l'exposition, ces valeurs peuvent influencer notablement la moyenne et, par voie de conséquence, fausser la lecture de la tendance centrale de la distribution de données. La médiane et la moyenne géométrique sont moins sensibles aux valeurs élevées. Le choix de la mesure dépendra des objectifs de l'étude.

Mesures descriptives d'une distribution asymétrique



### Test d'hypothèse

Comparaison de populations par une application rigoureuse de la théorie statistique de la décision. Selon l'approche classique (Neyman-Pearson), on formule une *hypothèse nulle*, que l'on teste par rapport à une *hypothèse alternative* au moyen d'une *variable à tester*. L'hypothèse nulle correspond toujours à l'absence de différence ou à l'absence d'effet, tandis que l'hypothèse alternative correspond à tous les autres scénarios.

*Exemple* : L'hypothèse nulle permettant de tester si les moyennes de deux populations,  $\mu_1$  et  $\mu_2$ , sont égales est formulée comme suit :

$$H_0: \mu_1 = \mu_2$$

et l'hypothèse alternative s'écrit comme suit :

$$H_1: \mu_1 \neq \mu_2$$

Il s'agit d'un test bilatéral. Dans le cas d'un test unilatéral, le signe  $\neq$  serait remplacé par  $<$  ou  $>$ , selon le cas.

### Observation des effets statistiquement significatifs

#### Comparaison de groupes

Les méthodes servant à établir la signification statistique des effets dépendent du type de données.

### *Données continues*

Les tests statistiques reposent sur un certain nombre d'hypothèses concernant les données. Il est important de vérifier ces hypothèses afin de garantir la validité du test statistique. Supposons que 10 animaux ont été exposés à une substance et qu'il existe 4 niveaux d'exposition. Nous cherchons à déterminer s'il y a eu une variation de poids perceptible, le poids étant en l'occurrence une variable continue. On se sert habituellement de l'analyse de variance pour déterminer s'il existe une différence statistiquement significative entre les moyennes des quatre groupes qui correspondent aux niveaux d'exposition. Or, nous devons formuler trois hypothèses au sujet des données pour garantir la validité de l'analyse de variance :

- (1) les données sont indépendantes les unes des autres;
- (2) les observations ont la même variance;
- (3) les poids obéissent à une distribution normale.

L'indépendance des données est généralement réalisée si la répartition des animaux entre les traitements se fait selon un processus de randomisation rigoureux. La normalité devrait être une caractéristique des données. On peut le vérifier de plusieurs manières. Si les données proviennent de la même distribution normale, l'histogramme représentant ces données aura un aspect symétrique, en forme de cloche. On croit souvent, à tort, que la représentation des données brutes d'une expérience devrait avoir cette forme si ces données proviennent d'une distribution normale. Ce n'est pas nécessairement vrai, s'il

existe des différences entre les moyennes des groupes. Si tel est le cas, l'histogramme pourrait avoir une forme irrégulière, indiquant ainsi que les données ont plus d'un mode. Pour vérifier la normalité, on peut soustraire la moyenne du groupe de chacune des valeurs de ce groupe. Si les données proviennent d'une distribution normale, l'histogramme représentant les *résiduels* devrait être en cloche, avec un seul mode.

<b>Normalité des résiduels</b>	
•	<b>Histogramme en cloche, non asymétrique</b>
•	<b>Variabilité uniforme dans le graphique du rapport entre les résiduels et les valeurs prédites</b>
•	<b>Test de normalité ne doit pas être significatif</b>

Parfois, après correction des moyennes, la courbe des résiduels prend un aspect asymétrique, avec une longue queue en pointe. On observe cela habituellement pour les données de concentration. Puisque l'analyse de variance exige la normalité, on doit "normaliser" les données au moyen d'une transformation avant d'effectuer le test. Le plus souvent, on appliquera une transformation logarithmique aux données de concentration pour les normaliser. D'autres types de données pourront exiger un autre mode de transformation. Une fois la normalité obtenue, on applique les tests statistiques au moyen des données transformées. Si l'échantillon est grand, obtenir la normalité est moins essentiel, puisque la distribution des moyennes tend vers la normalité lorsque la taille de l'échantillon augmente.

Si les hypothèses mentionnées ci-dessus sont applicables, on utilise le test t de Student pour tester l'hypothèse que les moyennes de deux groupes sont égales. Lorsqu'il y a plus de deux groupes, on recourt à l'analyse de variance.

<b>Tests de signification pour l'égalité de moyennes avec données normales</b>		
<b>Test</b>	<b>Description</b>	<b>Hypothèses</b>
<b>Test t</b>	<b>Égalité des moyennes de deux groupes de données indépendants</b>	<ul style="list-style-type: none"> <li>• <b>normalité</b></li> <li>• <b>indépendance de toutes les données</b></li> <li>• <b>variance homogène</b></li> </ul>
<b>Test t de comparaison par paires</b>	<b>Égalité des moyennes de groupes de données appariés (l'équivalent du test t à un seul échantillon pour les différences entre les paires de données)</b>	<ul style="list-style-type: none"> <li>• <b>normalité</b></li> <li>• <b>indépendance de toutes les paires de données</b></li> <li>• <b>variance homogène</b></li> </ul>
<b>Analyse de variance</b>	<b>Égalité des moyennes de plus de deux groupes de données indépendants</b>	<ul style="list-style-type: none"> <li>• <b>normalité</b></li> <li>• <b>indépendance de toutes les données</b></li> <li>• <b>variance homogène</b></li> </ul>

L'analyse de variance permet d'établir que au moins deux moyennes diffèrent significativement l'une de l'autre, mais elle n'indique pas quelles moyennes diffèrent les unes des autres. Pour connaître ces moyennes, on fait des tests de comparaisons multiples. Ce n'est pas aussi simple qu'appliquer des tests t de comparaison par paires au groupe témoin et aux groupes expérimentaux, puisque les niveaux de signification des tests d'hypothèses ne sont plus valides, d'où la possibilité de recourir à une méthode de comparaison multiple. Le test de Dunnett convient très bien pour la comparaison du groupe témoin avec chacun des autres groupes. Le test de Duncan (Multiple Range test), qui est une autre méthode de comparaison multiple, se prête bien à des comparaisons par paires de toutes les paires possibles.

<b>Comparaisons multiples</b>		
<b>Test de Dunnet</b>	<b>Comparaison des moyennes de plusieurs groupes à celle du groupe témoin</b>	<ul style="list-style-type: none"> <li>• normalité</li> <li>• indépendance</li> <li>• variances égales</li> </ul>
<b>Test de Duncan</b>	<b>Comparaison de toutes les paires de traitements possibles</b>	<ul style="list-style-type: none"> <li>• normalité</li> <li>• indépendance</li> <li>• variances égales</li> <li>• groupes de même taille</li> </ul>

### *Données catégorielles*

On parle de données catégorielles lorsque la valeur observée peut faire l'objet d'un classement parmi un petit nombre de catégories. Par exemple, dans une étude de l'action carcinogène, les animaux sont classés soit comme porteur d'une tumeur ou comme non porteur. On peut ensuite déterminer l'effectif de chaque catégorie. Voici un exemple de données catégorielles.

<b>Données binomiales</b>		
<b>Groupe</b>	<b>Nombre d'animaux porteurs d'une tumeur</b>	<b>Nombre total d'animaux</b>
<b>Témoin</b>	<b>1</b>	<b>100</b>
<b>Expérimental</b>	<b>4</b>	<b>50</b>

Notons que le nombre d'animaux non porteurs d'une tumeur s'obtient en soustrayant le

nombre d'animaux porteurs du nombre total d'animaux. Le test exact de Fisher sert à comparer la proportion d'animaux porteurs d'une tumeur dans le groupe témoin et la proportion d'animaux porteurs dans le groupe expérimental.

Il y a des études où les résultats peuvent être classés dans plus de deux catégories, par exemple les tests d'irritation des yeux ou de la peau, où l'on cote le degré d'irritation. Les études de ce genre exigent une méthode d'analyse plus perfectionnée. Le programme SAS contient l'application CATMOD, qui peut exécuter un large éventail d'analyses de données catégorielles (SAS, 1990). Dans Shoukri et Edge (1996), on fait une bonne analyse de quelques-unes de ces techniques.

#### *Données de dénombrement*

Les données de dénombrement seront utiles lorsque le nombre d'observations (ou de résultats) a de l'importance et que ce nombre n'est pas assujéti à un maximum, par exemple le nombre de mutations dans une bio-analyse. Le différentiel de leucocytémie et le nombre de globules rouges nucléés sont d'autres exemples de données de dénombrement. Les données de ce type suivent une distribution de Poisson. Une façon simple d'analyser des données de dénombrement est d'extraire la racine carré de chaque observation, puis d'appliquer les tests habituels, comme il en est fait mention dans la section 2.2. Cette méthode donne de très bons résultats lorsque les chiffres sont relativement élevés. Si les chiffres sont petits, il sera plus utile de recourir à des méthodes plus perfectionnées, qui font intervenir la distribution de Poisson pour tenir compte de la dispersion imputable à l'erreur expérimentale.

#### *Tests non paramétriques*

On appelle *test paramétrique* le test fondé sur la normalité qui s'applique aux données continues, parce qu'on suppose que la distribution d'où proviennent les données est connue, en l'occurrence la distribution normale. Or, il arrive que l'on ne connaisse pas la distribution d'où sont tirées les données. Par exemple, on pourrait ne pas trouver de formule de transformation adéquate pour normaliser des données continues. La seule hypothèse raisonnable serait que la forme de la distribution est la même pour chaque groupe étudié, mais que la position peut varier. Dans ce cas, on utilisera des tests paramétriques pour déceler les changements de position significatifs.

Le test de Wilcoxon (somme des rangs) et le test de Mann-Whitney (test U) servent à déterminer si deux groupes ont les mêmes paramètres de position, tandis que le test de Kruskal-Wallis vise à déterminer si  $k$  groupes ( $k > 2$ ) sont tous identiques ou si au moins un des groupes tend à fournir des observations différentes des autres groupes. Ces tests reposent tous sur le rang des observations plutôt que sur les observations proprement dites.

### ***Analyse de régression***

Lorsque la variable étudiée est continue, par exemple poids ou concentration, et que l'on s'intéresse au rapport entre cette variable et une autre variable continue, par exemple la taille, on utilisera l'analyse de régression pour déterminer si ce rapport est significatif et de quel type de rapport il s'agit. La formule de régression linéaire simple avec une seule variable indépendante s'écrit comme suit :

$$y = \alpha + \beta x + \varepsilon$$

où  $y$  est la variable étudiée,  $x$ , la variable indépendante,  $\alpha$ , le paramètre d'ordonnée à

l'origine,  $\beta$ , la pente, et  $\varepsilon$ , un terme d'erreur qui, généralement, est censé suivre une distribution normale de moyenne nulle et de variance inconnue. La pente peut être croissante, décroissante ou fixe. L'ordonnée à l'origine est la valeur moyenne de  $y$  lorsque  $x$  est égal à 0.

Il est possible d'élargir la régression linéaire simple en  $y$  ajoutant des variables indépendantes (ou explicatives) : c'est ce qu'on appelle la régression multiple. Celle-ci peut comprendre aussi des variables indépendantes qualitatives. On appelle modèle linéaire général la combinaison de l'analyse de variance et de l'analyse de régression. Une des conditions s'appliquant aux variables explicatives est que celles-ci doivent être indépendantes les unes des autres. Sinon, il y a multicollinéarité et les conclusions de l'analyse risquent d'être ambiguës.

Les commentaires précédents supposent que la variable étudiée dépend linéairement des variables indépendantes. Si ce n'est pas le cas, il faut recourir à la régression non linéaire. Par exemple :

$$y = \alpha \exp(-\beta x)\varepsilon$$

est une expression non linéaire courante. L'analyse de régression non linéaire se fait normalement au moyen d'un progiciel d'analyse statistique. Toutefois, dans le cas de l'expression ci-dessus, on pourrait se servir de la régression linéaire si on appliquait une transformation logarithmique au modèle, ce qui donnerait :

$$\log y = \log \alpha - \beta x + \log \varepsilon.$$

L'analyse porterait alors sur la nouvelle variable indépendante  $\log y$  et serait plus simple

que la régression non linéaire, technique qui exige une connaissance poussée des méthodes statistiques. La difficulté que pose l'analyse de données transformées a trait à l'interprétation des modèles estimés.

Là aussi, les hypothèses concernant les données doivent se vérifier avant l'exécution de l'analyse. On peut vérifier les critères de linéarité, de normalité, d'indépendance et d'homoscédasticité pour les résiduels de l'analyse. Neter *et al.* (1990) analysent en profondeur de nombreux aspects des modèles linéaires généraux.

On peut aussi appliquer l'analyse de régression à des variables qualitatives catégorielles. La technique appelée régression logistique est décrite plus en détail dans Shoukri et Edge (1996). Le programme SAS contient une application d'analyse logistique (SAS, 1990).

### ***Analyse multivariée***

On peut recourir à l'analyse multivariée lorsqu'un problème suppose plusieurs variables dépendantes et que l'étude a pour objectif de définir non seulement le rapport entre les variables indépendantes et les variables dépendantes, mais celui entre les variables étudiées. Les méthodes d'analyse multivariée les plus courantes sont l'analyse en composantes principales, l'analyse typologique, l'analyse factorielle et l'analyse multidimensionnelle. On trouvera un examen complet de la question dans Johnson et Wichern (1992). Les programmes informatiques tels que SAS sont particulièrement utiles pour l'analyse multivariée (SAS, 1990).

## **SUJETS PARTICULIERS**

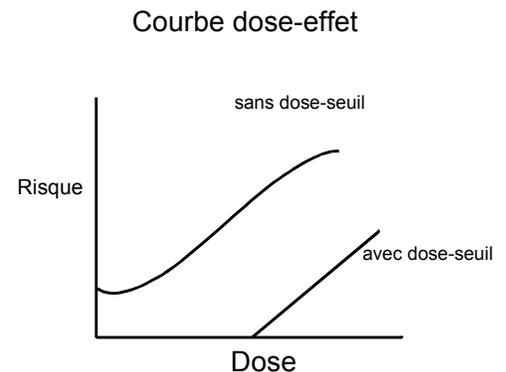
### **Estimation de l'activité**

### ***Toxiques avec dose-seuil***

En matière de santé, on définit des lignes directrices et on établit des priorités en se fondant sur les estimations de l'activité des substances toxiques. Deux types de mécanismes existent pour déterminer quelle méthode utiliser pour estimer l'activité des toxiques. Pour certains effets, on croit qu'il existe un seuil d'exposition en-deçà duquel l'organisme ne subit aucune modification pathologique. Pour d'autres, comme les effets génotoxiques (substances carcinogènes), on croit qu'il n'existe aucun seuil.

La figure ci-contre décrit la relation dose-effet théorique pour les toxiques avec dose-seuil et ceux sans dose-seuil.

La notion de seuil repose sur l'hypothèse que le niveau posologique détermine si l'exposition entraînera ou non un effet pathologique. Idéalement, du point de vue de la gestion du risque, il faudrait estimer ce seuil. Or, il peut varier d'un individu à l'autre. Selon la méthode d'évaluation du risque basée sur le facteur d'incertitude, pour les toxiques avec dose-seuil, on établit la *dose qui n'entraîne pas d'effet observé* (DSEO) en se fondant sur des expériences dose-effet menées sur des animaux. En raison des incertitudes inhérentes à l'extrapolation inter-espèces de l'animal à l'homme, et en raison de la variabilité de la sensibilité humaine aux effets des toxiques, on introduit un facteur de sécurité, ou facteur d'incertitude (FI). Ainsi, on pourra calculer, par exemple, la dose journalière admissible (DJA) en divisant la valeur de la DSEO par le facteur d'incertitude FI, c'est-à-dire :



$$\text{DJA} = \text{DSEO} / \text{FI}.$$

La valeur des facteurs d'incertitude varie de 1 à 10, 10 étant la valeur implicite la plus courante. Il existe d'autres facteurs de sécurité pour tenir compte de l'utilisation de la DMEDO au lieu de la DSEDO et compenser les irrégularités de la base de données.

La méthode fondée sur le facteur d'incertitude fait l'objet d'un certain nombre de critiques. Premièrement, la dose sans effet observé (DSEO) dépend de la taille de l'échantillon. Par exemple, un taux de réponse de 0 sur 10 n'a pas la même signification qu'un taux de 0 sur 1 000. En outre, cette méthode suppose l'existence d'une dose-seuil en-deçà de laquelle l'organisme ne subit aucune modification pathologique. Ce seuil peut varier d'un animal à l'autre, en sorte qu'il faudrait déterminer un niveau minimum.

Les spécialistes sont de plus en plus nombreux à penser que la méthode de la dose-repère pourrait résoudre certaines de ces difficultés. Cette méthode consiste à modéliser la courbe dose-effet, puis à calculer la limite inférieure de l'intervalle de confiance à 95 % relatif à la dose qui correspond à une faible augmentation du risque par rapport au taux de référence.

On peut appliquer la méthode de la dose-repère à des variables discrètes (réponse par tout ou rien) comme à des variables continues.

Type de données	Modèle	Équation du modèle
Discrètes	Régression linéaire	$P(d)=c+(1-c\{1-\exp[-q_1(d-d_0)]\})$
	À plusieurs degrés	$P(d)=1+\exp[-q_1(d-d_0)-q_2(d-d_0)^2- \dots -q_k(d-d_0)^k]$

Continues	Régression linéaire	$m(d)=c+q_1(d-d_0)$
	Régression polynomiale	$m(d)=c+q_1(d-d_0)+q_2(d-d_0)^2+ \dots +q_k(d-d_0)^k$

Il existe d'autres méthodes d'ajustement de courbe.

En ce qui concerne les données discrètes, la dose-repère  $d$  est la dose qui crée un risque additionnel déterminé, défini par l'expression

$$\frac{(P(d) - P(0))}{(1 - P(0))}$$

où  $P(0)$  est la valeur de l'équation du modèle,  $P(d)$ , lorsque la dose est nulle. Dans le cas des données continues, la dose-repère est la dose qui produit l'effet additionnel spécifié :

$$\frac{m(d) - m(0)}{m(0)}$$

On introduit alors un facteur de sécurité pour ramener la dose-repère préétablie de 1 - 10 % à un niveau de risque "acceptable". Pour une analyse plus complète de la dose-repère, voir McColl (1990, p. 23).

### ***Toxiques sans dose-seuil***

En ce qui concerne les toxiques sans dose-seuil, on suppose que toute exposition peut entraîner des effets toxiques, de sorte qu'il est possible de modéliser la relation exposition-effet à l'aide d'une courbe lisse qui tend vers les taux de référence lorsque le niveau

d'exposition diminue. La question maintenant est donc de savoir comment exprimer le risque.

En ce qui concerne l'évaluation des substances d'intérêt prioritaire aux termes de la Loi canadienne sur la protection de l'environnement, l'évaluation quantitative du risque avait pour objectif de produire un classement des substances carcinogènes. À cette fin, on a calculé directement à partir de la courbe exposition-effet l'exposition à vie qui implique une hausse de 5 % du risque. On a combiné cette valeur à l'exposition estimée de la population canadienne en général pour obtenir l'indice du risque d'exposition pour chaque substance. Les indices ainsi obtenus pouvaient ensuite servir à l'opération de classement.

D'autres programmes, comme ceux ayant trait aux déchets dangereux ou à l'eau potable, visent surtout à la définition de principes directeurs. Dans ce cas, les responsables de l'action gouvernementale doivent à tout prix disposer de suffisamment d'informations pour pouvoir envisager des mesures correctives pour l'avenir. En règle générale, la courbe dose-effet produit une estimation du risque excédentaire à des valeurs de dose peu élevées.

Pour l'un et l'autre objectifs, on estime le risque en modélisant les effets des toxiques en fonction de l'exposition. Krewski et VanRyzin (1981) analysent et comparent un certain nombre de modèles appartenant à l'une ou l'autre de deux catégories. Les modèles statistiques, ou de tolérance, reposent sur l'hypothèse qu'un individu échantillonné au hasard présentera un effet à la date  $d$  selon la probabilité suivante :

$$P(d) = \Pr(\text{tolérance} \leq d) = F(a + b \log(t)).$$

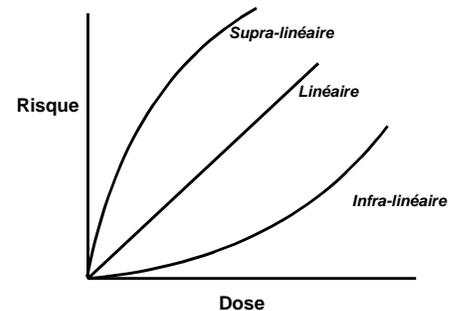
Cette catégorie de modèles comprend le modèle probit, le modèle logistique et le modèle

des valeurs extrêmes.

Les modèles stochastiques ou mécanistiques reposent, quant à eux, sur l'hypothèse qu'un effet est induit par suite de la survenue aléatoire d'un ou de plusieurs événements biologiques. Par exemple, dans le cas du modèle de l'exposition unique, un effet sera induit après que l'organe-cible aura reçu une unité de dose biologiquement efficace dans un intervalle donné. Dans le cas du modèle des expositions multiples successives, on suppose qu'un effet pourra être induit après plusieurs expositions. Quant au modèle à plusieurs degrés, il suppose que des effets toxiques auto-reproducteurs irréversibles seront induits par la survenue de divers événements biologiques aléatoires et que, en l'occurrence, la fréquence par âge de chaque événement dépendra linéairement de la dose.

D'autres programmes de réglementation ont pour objet spécifique de déterminer le risque de cancer aux niveaux d'exposition ambiants. Les niveaux ambiants sont ordinairement beaucoup moins élevés que les niveaux d'exposition appliqués dans les tests biologiques faits sur des animaux. Comme les modèles exposition-effet sont établis en fonction de ces niveaux d'exposition relativement élevés, il faut faire une extrapolation des risques correspondant aux niveaux plus élevés pour connaître les risques correspondant aux niveaux plus bas. Toutefois, ces modèles sont sensibles aux groupes expérimentaux, puisqu'on ne connaît pas la forme que prend la courbe pour les valeurs les plus basses.

Courbes dose-effet pour de faibles doses



### *Modèle multi-degrés*

Le modèle multi-degrés est employé par l'EPA aux États-Unis et d'autres organismes de réglementation pour estimer le risque de cancer. Il est fondé sur la probabilité  $P(d)$  qu'une tumeur se forme par suite de l'exposition d'un sujet à une dose  $d$  à une période déterminée  $t$ , soit :

$$P(d) = 1 - \exp[-ct^k \sum_{i=0}^k (q_i d^i)]$$

où  $k$  est le nombre hypothétique de degrés (ou stades) dans l'action cancérigène. À mesure que la dose diminue, la relation entre le risque excédentaire et la dose se rapproche de la linéarité. On calcule alors la limite supérieure de l'intervalle de confiance à 95 %, et cette limite devient le facteur pente pour les valeurs de dose peu élevées. L'inverse de la pente s'appelle le risque unitaire. On appelle cette technique la méthode multi-degrés linéarisée (*Linearized Multistage Method – LMS*).

#### *Modélisation linéaire robuste*

Le fait d'extrapoler à partir de modèles alignés sur des niveaux d'exposition relativement élevés pour déterminer le risque aux niveaux d'exposition ambiants pose un certain nombre de difficultés que l'on n'a pas manqué de souligner. Krewski *et al.* (1991) ont élaboré une méthode (appelée *model free extrapolation method – MFX*, ou *robust linear modelling*) qui résout certaines de ces difficultés. Cette méthode nécessite une seule hypothèse, à savoir que la courbe dose-effet est linéaire pour les valeurs de dose peu élevées. Considérons un test biologique impliquant  $t+1$  groupes,  $y$  compris le groupe-témoin. Calculons la probabilité d'une réponse dans chaque groupe au moyen de la formule  $p_i = r_i/n_i$ , où  $r_i$  est le nombre d'animaux du groupe  $i$  sur lesquels on observe des

effets et  $n_i$ , l'effectif du groupe  $i$ . Comparons la valeur  $p$  pour le groupe-témoin à la valeur  $p$  de chaque groupe en utilisant le test exact de Fisher à un seuil de signification de 5 %. Ensuite, calculons la limite supérieure de l'intervalle de confiance à 95 % pour la valeur  $p$  de chaque groupe expérimental qui précède le premier groupe significatif, ainsi que la limite inférieure de l'intervalle de confiance à 95 % pour la valeur  $p$  du groupe-témoin, en utilisant les intervalles binomiaux exacts et le facteur  $5/(t+1)$  comme niveau de signification selon l'inégalité de Bonferroni. Enfin, calculons la pente des droites qui relient ces limites supérieures et la limite inférieure calculée pour le groupe-témoin et choisissons la pente la moins élevée comme facteur de risque.

Les principaux avantages de la méthode MFX sont les suivants : i) on ne suppose aucun modèle qui puisse influencer notablement sur l'extrapolation pour les faibles doses, et ii) il est raisonnablement permis de croire que le risque réel est inférieur au risque calculé au moyen de la méthode MFX. Parmi les faiblesses de la méthode, notons le fait que les données ne sont pas toutes utilisées pour l'estimation de l'activité à de faibles doses. Krewski *et al.* (1991) montrent que le rapport médian de l'estimation MFX de la pente pour les faibles doses à l'estimation LMS correspondante est égal à 1,3. Donc, les valeurs de pente calculées à l'aide de la méthode MFX sont légèrement supérieures à celles calculées avec LMS. Dans 443 des 572 essais, le rapport était inférieur à 2.

Le calcul du risque de cancer par les méthodes standard repose sur un protocole d'exposition qui peut être appliqué par intermittence, mais toujours de la même manière, durant toute la vie de l'animal. Pour connaître le risque correspondant à une dose journalière moyenne fixe exprimée en mg/kg pv (poids vif en grammes)/jour, il faut apporter une correction à l'exposition. Par exemple, dans une expérience d'inhalation des souris ont été exposées à des concentrations dans l'atmosphère 8 heures par jour, 5 jours par semaine, pendant 92 semaines. La durée de vie normale d'une souris est 104 semaines.

Pour convertir la durée d'exposition réelle en durée d'exposition vie entière, on pondère la concentration en laboratoire de la manière suivante :

$$\text{conc. en exposition constante} = \text{conc. expérimentale} \times (5/7) \times (8/24) \times (92/104).$$

Il arrive que la durée de l'expérience ne soit pas aussi longue que la durée de vie normale de l'animal étudié. Si l'expérience cesse relativement tôt dans la vie de l'animal, on dénumbrera probablement moins d'animaux porteurs d'une tumeur que si l'expérience s'était étalée sur toute la durée de vie. Dans des comparaisons avec d'autres cancérigènes, on observe que l'activité semble moins forte. Des essais permettent de croire que le taux de formation des tumeurs s'accroît non linéairement avec l'âge (la relation étant plutôt exponentielle); c'est pourquoi il convient de redresser le taux de formation des tumeurs par un facteur  $f^2$  ou  $f^3$ , où :

$$f = \text{durée de l'essai sur l'animal} / \text{durée de vie normale}.$$

Il arrive aussi que l'exposition à une substance ne soit pas constante ni uniforme dans le temps. Par exemple, il se peut que les enfants et les petits enfants soient exposés à de plus fortes doses de résidus de pesticides dans le jus de pomme (en mg/kg pv/jour) que les adultes. Murdoch, Krewski et Wargo (1992) examinent cette question par rapport aux modèles de mutation naissances-décès à deux degrés et à degrés multiples. Ils observent que l'utilisation de la dose journalière moyenne à vie comme facteur d'approximation du risque à vie peut mener à la surestimation ou à la sous-estimation du risque, suivant les modalités d'exposition.

### ***Extrapolation de l'animal à l'homme***

Les animaux utilisés dans les expériences de toxicologie sont de plus petite taille que les humains et leur taux de métabolisme est habituellement plus élevé. C'est pourquoi la surface corporelle sert souvent de base à l'extrapolation, parce que, croit-on, elle tient mieux compte de la variation de taux de métabolisme d'une espèce à une autre que le poids corporel. Il est alors admis que différentes espèces ont la même sensibilité à une dose donnée par unité de surface corporelle. Le facteur de conversion le plus courant dans les circonstances est :

$$\text{facteur de correction inter-espèces} = 3(70/w)^{1/2}$$

où 70 (kilogrammes) est le poids corporel type d'un humain et w (kilogrammes), le poids corporel type d'un animal, étant donné l'espèce utilisée pour l'étude.

D'autres facteurs de conversion sont proposés, et on peut les exprimer suivant la formule générale suivante :

$$\text{activité} = a \times (\text{poids corporel})^b$$

où b peut prendre différentes valeurs (voir tableau ci-dessous).

<b>b</b>	<b>Extrapolation inter-espèces fondée sur</b>
<b>1</b>	<b>poids corporel</b>
<b>2/3</b>	<b>surface corporelle</b>
<b>3/4</b>	<b>taux de métabolisme</b>

## 4.2 Évaluation probabiliste

L'évaluation de l'exposition et du risque dans les études d'impact sur la santé comporte de nombreux éléments d'incertitude. Par exemple, on estime souvent le risque à partir d'études de toxicologie faites sur des animaux. Ces études sont conduites selon des protocoles rigoureux sur des populations homogènes de rongeurs. Évaluer le risque par l'extrapolation du rongeur à l'homme est une opération remplie d'incertitude. En l'absence de données mécanistiques prouvant le contraire, on suppose que les humains courent le même risque que les rongeurs. Néanmoins, un certain nombre d'études confirment que ce n'est pas toujours le cas; ces études montrent que les effets observés chez certaines espèces de rongeurs ne se retrouvent pas chez d'autres espèces.

Parmi les autres facteurs d'incertitude notons l'extrapolation des risques correspondant aux doses élevées utilisées en laboratoire pour connaître les risques correspondant aux doses plus faibles reçues par les humains (milieu ambiant); l'extrapolation des résultats d'un essai toxicologique obtenus selon une modalité d'exposition déterminée (voie orale, inhalation, voie de pénétration cutanée) en fonction d'une autre modalité d'exposition; ou encore, l'utilisation des résultats d'études subchroniques en vue d'estimer des seuils pour des études dose-effet chroniques. Les données épidémiologiques concernant des populations humaines exposées peuvent elles aussi être marquées fortement par l'incertitude, notamment lorsqu'il est difficile d'établir les profils d'exposition rétrospectifs, particulièrement dans le cas de maladies chroniques comme le cancer, pour lesquelles les données sur l'exposition doivent remonter à de nombreuses années avant le moment où la maladie a été diagnostiquée. Même les études prospectives comportent un élément d'incertitude appréciable. Par exemple, dans les études alimentaires, la mémorisation et le journal alimentaire ne sont pas à l'abri des erreurs. Il en est de même pour les diagnostics.

---

La façon classique d'évaluer le risque consiste à calculer des estimations ponctuelles du risque en s'appuyant sur des hypothèses prudentes. Or, les estimations ponctuelles reposent sur des données très incertaines. En focalisant notre attention sur des estimations globales uniques, nous risquons d'oublier que ces estimations s'assimilent plus à une limite supérieure qu'à la meilleure valeur estimée.

En caractérisant le risque d'une manière plus complète par l'analyse d'incertitude, nous pouvons mieux comprendre le rôle de l'incertitude dans les estimations du risque et de l'exposition. Nous pouvons résumer les résultats sous la forme d'une distribution des niveaux de risque ou des niveaux d'exposition possibles afin de tenir compte du plus grand nombre de facteurs d'incertitude ou de variabilité possible. Pour une description plus détaillée de l'analyse d'incertitude, voir Morgan, Henrion et Small (1990), NRC (1994) et Bartlett *et al.* (1996).

Dans un exemple, Hoffman et Hammonds (1994) utilisent un modèle de risque multiplicatif pour estimer le rapport de risques pour une population d'individus ayant mangé du poisson contaminé. Le risque  $R$  est défini par l'expression :

$$R = X_1 \times X_2 \times (X_3)^{-1} \times (X_4)^{-1}$$

où  $X_1$  est le taux de concentration du contaminant dans le poisson,  $X_2$ , la quantité de poisson introduite dans l'estomac,  $X_3$ , la masse corporelle d'un humain, et  $X_4$ , la dose de référence. On établit la distribution d'incertitude globale pour  $R$  tout d'abord en estimant les distributions des variables indépendantes ( $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$ ), en échantillonnant des valeurs de chacune de ces variables au moyen d'une analyse de Monte Carlo, en intégrant les valeurs échantillonnées dans l'équation du risque, puis en calculant  $R$ . On répète cette

opération des milliers de fois afin de définir une distribution d'échantillonnage pour  $R$ . Bien que des programmes comme Crystal Ball et @Risk rendent facile et commode l'analyse de Monte Carlo, celle-ci doit être exécutée avec soin pour garantir l'intégrité des résultats (Burmester et Anderson, 1994).

Rai *et al.* (1996) soulignent l'importance de savoir faire la distinction entre la variabilité propre aux paramètres de l'étude et l'incertitude attribuable à un manque de connaissances. Dans l'exemple précédent, les auteurs supposent que  $X_4$  comporte uniquement un élément d'incertitude,  $X_3$  comporte uniquement un facteur de variabilité, et  $X_1$  et  $X_2$  comportent les deux. On construit des distributions pour chaque composante, puis on les utilise dans une analyse de Monte Carlo. On peut alors évaluer l'effet des divers facteurs de variabilité et d'incertitude individuellement ou globalement.

### **Contraintes liées à l'utilisation de la biostatistique**

Lorsqu'on utilise la biostatistique dans le but d'appuyer la prise de décisions par des données fiables, on doit faire face à un certain nombre de contraintes.

- La complexité de l'analyse peut exiger une excellente compréhension des concepts statistiques, si l'on veut pouvoir utiliser ou réviser cette analyse. Plus le protocole d'acquisition des données est complexe, plus les méthodes statistiques nécessaires pour l'analyse seront complexes.
- Il est toujours possible d'appliquer des méthodes statistiques simples, mais elles doivent être interprétées en fonction du mode de collecte des données.
- On peut devenir plus à l'aise avec les méthodes statistiques choisies en recevant

---

une formation appropriée ou en consultant des statisticiens d'expérience, le cas échéant. Il n'est pas toujours nécessaire de consulter un statisticien; il importe plus de maîtriser suffisamment les concepts, de manière que l'on sache à quel moment il faut recourir aux services d'un statisticien.

- Idéalement, on doit consulter un statisticien qui connaît très bien les méthodes utilisées dans les études d'impact sur la santé. Par exemple, les facteurs de risque sont élaborés à l'aide de méthodes spécifiques. Il est essentiel de connaître les limites des données sur lesquelles reposent les facteurs de risque pour pouvoir bien interpréter ces facteurs.

## Conclusions

- Les méthodes statistiques peuvent accroître la crédibilité des résultats.
- Elles fournissent aussi une bonne indication de la variabilité des données.
- On doit choisir soigneusement les méthodes d'analyse et interpréter aussi soigneusement les résultats de l'analyse.
- Il est utile de savoir à quel moment consulter un statisticien. L'idéal est de pouvoir compter sur un statisticien qui a de l'expérience dans les études d'impact sur la santé ou les évaluations environnementales.

## Ressources documentaires, spécialistes, organisations

Il existe de nombreux bons ouvrages où l'on explique en détail les notions

*Outils d'apprentissage :*  
Manuels de base pour non-spécialistes, cours d'introduction, CD-ROM multimédia.

élémentaires de statistique. Il est préférable de choisir des manuels rédigés pour des non-spécialistes, car ceux destinés aux statisticiens se caractérisent plutôt par un vocabulaire et une présentation très techniques. Le non-spécialiste sera bien avisé de suivre un cours de statistique avant d'entreprendre de l'auto-formation. La statistique repose sur des concepts de probabilité abstraits qu'il est plus facile d'apprendre dans un environnement interactif.

Les cours et les manuels peuvent donner l'essentiel des concepts statistiques, mais ils ne peuvent couvrir tout le domaine de la statistique, tellement celui-ci est vaste. Comme nous l'avons mentionné plus haut, les méthodes statistiques reposent toujours sur des hypothèses concernant les données. Les statisticiens sont formés pour identifier ces hypothèses et choisir les méthodes pertinentes, ou encore élaborer ou modifier des méthodes, selon les besoins. Si l'on ne tient pas compte des hypothèses, les conclusions de l'analyse pourraient être erronées. Cela peut aller de l'erreur sans conséquence réelle à l'erreur très grave, qui peut, par exemple, entraîner des décisions coûteuses concernant l'atténuation des impacts, alors qu'une telle mesure n'aurait pas été justifiée si l'analyse des données s'était faite correctement.

Vous pouvez éviter des erreurs aussi coûteuses en consultant au moment opportun un statisticien au sein de votre organisation. Si celle-ci ne compte aucun statisticien dans ses rangs, vous pourrez vous tourner vers l'université de la région ou un bureau d'experts-conseils. Idéalement, le statisticien que vous consulterez devrait avoir une connaissance spécialisée des types de données et des questions liés aux études d'impact sur la santé. Par exemple, il devrait avoir de l'expérience dans la modélisation de l'exposition ou la conception d'études de toxicologie.

Il existe par ailleurs de nombreux programmes informatiques qui permettent de faire des

analyses statistiques. Des tableurs comme Lotus 1-2-3, Excel et Quattropro peuvent faire de l'analyse de régression ou dessiner des graphiques et ils comportent d'autres fonctions statistiques en plus de leur fonctions de chiffrier. D'autres programmes sont conçus spécialement pour l'analyse statistique (p. ex., SAS, SPlus, SPSS, etc.). Des programmes complémentaires, comme Crystal Ball (Decisioneering, 1996) exécutent des analyses de Monte Carlo élémentaires. Enfin, des CD-ROM multimédia interactifs offrent des cours de statistique, p. ex. ActivStats (Velleman, 1998).

**Bibliographie**

Bartlett, S., Richardson, G.M., Krewski, D., Rai, S.N., et Fyfe, M. (1996) "Characterizing Uncertainty in Risk Assessment - Conclusions Drawn from a Workshop", *Human and Ecological Risk Assessment*, 2, pp. 221-231.

Burmester, D.E. et Anderson, P.D. (1994) "Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments", *Risk Anal.*, 14, pp. 477-481.

Decisioneering (1996) *Crystal Ball, Version 4.0 User Manual*. Denver, Colorado, Decisioneering Inc.

Hoffman, F.O. et Hammonds, J.S. (1994) "Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability", *Risk Anal.*, 14, pp. 707-712.

Johnson, R.A. et Wichern, D.W. (1992) *Applied Multivariate Statistical Analysis*. 3<sup>rd</sup> ed., Englewood Cliffs, New Jersey, Prentice-Hall Inc., 642 pages.

Kennedy, John B. et Neville, Adam M. (1974) *Basic Statistical Methods for Engineers & Scientists*, 2<sup>nd</sup> ed., New York, Harper & Row Publishers, 490 pages.

Krewski, D. et Van Ryzin (1981) "Dose response models for quantal toxicity data", dans *Statistics and Related Topics*, M. Csorgo, D.A. Dawson, J.N.K. Rao, A.K.Md.E. Saleh (dir.), North-Holland Publishing Company, 1981, pp. 201-231.

Krewski, D., Gaylor, D., et Szyszkowicz (1991) "A model-free approach to low-dose extrapolation", *Environmental Health Perspectives*, **90**, pp. 279-285.

MathSoft (1997) *S-Plus 4 Guide to Statistics*. Seattle, Washington, MathSoft Inc.

McColl, R.S. (1990) "Facteurs de sécurité biologiques et évaluation du risque toxicologique", *Rapport numéro 90-DHM-154 de Santé Canada*.

Morgan, M., Henrion, M., et Small, M. (1990) *Uncertainty*. New York, Cambridge University Press, 332 pages.

Murdoch, D.J., Krewski, D. et Wargo, J. (1992) "Cancer Risk Assessment with Intermittent Exposure", *Risk Analysis*, **12**, pp. 569-577.

Nete, J., Wasserman, W., Kutner, M.H. (1990) *Applied Linear Statistical Models*. 3<sup>d</sup> ed. Homewood, IL, Richard D. Irwin Inc., 1181 pages.

Rai, S.N., Krewski, D., et Bartlett, S. (1996) "A General Framework for the Analysis of Uncertainty and Variability in Risk Assessment", *Human and Ecological Risk Assessment*, **2**, pp. 972-989.

Shoukri, M.M., Edge, V.L. (1996) *Statistical Methods for Health Sciences*, Boca Raton, CRC Press, 298 pages.

SPSS (1990) *SPSS Advanced Statistics Student Guide*, Chicago, Illinois, SPSS Inc.

Statistical Analysis System (SAS) (1990). *SAS/STAT User's Guide*, Cary, NC, SAS Institute Inc.

Velleman, P. (1998) *ActivStats 2.0*, Reading, MA, Addison Wesley Longman.