Analyse statistique des données dosimétriques d'exposition externe des travailleurs

J. BRENOT, P.G. BEAU, N. PARMENTIER, J. PIECHOWSKI

(Manuscrit reçu le 3 octobre 1985)

RÉSUMÉ

L'utilisation des données mensuelles d'exposition externe des travailleurs se situe à différents niveaux : l'individu isolé, l'individu dans son groupe, le groupe en soi, enfin les groupes entre eux.

Dans chaque cas, des indicateurs statistiques (dose annuelle, dose collective annuelle, dose annuelle moyenne, etc.) sont employés pour résumer une situation individuelle ou collective. La pertinence et l'efficacité des indicateurs sont discutées dans une perspective de gestion du risque radiologique individuel et collectif. La nécessité de recourir aux distributions statistiques apparaît dès qu'il faut répondre à un problème d'estimation, tant au plan individuel, c'est le cas des données manquantes, qu'au plan du groupe pour les doses annuelles des individus le composant ou pour les doses reçues par ceux-ci sur une certaine période qui peut être celle de la vie professionnelle.

ABSTRACT

The monthly monitoring data of workers' external exposures are used in different ways : to focus on the worker as an individual or as a group member, on the group as such, or to compare groups.

In each case, statistical dosimetric indicators are defined to summarize an individual status or a collective situation. The relevance and efficiency of these indicators are discussed in the context of radiological risk management. The use of statistical distributions cannot be avoided in situations requiring an estimation process, for instance, at the worker's level, missing data in his dose record, at the group level, modeling of annual doses and life-time dose distributions.

1. INTRODUCTION

La surveillance individuelle des travailleurs du groupe CEA soumis à une exposition externe est basée essentiellement sur l'exploitation de dosimètres photographiques qui intègrent l'exposition sur une période

* Commissariat à l'énergie atomique, Institut de protection et de sûreté nucléaire, Département de protection sanitaire, Services d'hygiène radiologique, BP 6, 92265 Fontenay-aux-Roses Cedex, France.

en général mensuelle ainsi que l'exige la législation. Le premier objectif de la surveillance est de vérifier que la dose pour chaque travailleur reste, chaque mois, inférieure à la valeur autorisée légalement. Le second est de pouvoir évaluer la dose individuelle dans le cadre de la vie professionnelle ; ceci implique le suivi, au cours du temps, des expositions individuelles au sein des groupes de travailleurs. Cette connaissance est indispensable pour une bonne gestion du système de radioprotection qui tienne compte des répartitions de doses dans les diverses sous-populations. Les indicateurs que sont, au niveau de l'individu, la dose mensuelle, la dose annuelle et la dose sur sa vie professionnelle ou dose-vie ainsi que leur distribution au niveau de sous-populations de travailleurs sont nécessaires pour atteindre ces objectifs.

Les problèmes statistiques qui se posent lors de l'exploitation des fichiers dosimétriques de relevés individuels mensuels, annuels ou de doses-vie seront abordés successivement.

2. LE FICHIER DES DOSES INDIVIDUELLES MENSUELLES

La donnée de base est constituée par le relevé dosimétrique mensuel fait, pour l'instant, à partir de l'exploitation d'un film dosimètre selon la législation française en vigueur. Deux facteurs entrent en jeu : la sensibilité du dosimètre qui est de l'ordre de 0,15 mSv et le seuil d'enregistrement admis habituellement et réglementairement égal à 0,25 mSv. De plus, pour tenir compte du fait que de nombreux travailleurs prennent leurs congés annuels pendant les mois de juillet et d'août, les résultats des dosimètres portés au cours de ces deux mois sont systématiquement regroupés ; chaque travailleur possède donc onze relevés dosimétriques mensuels par an.

2.1. Les pratiques classiques d'estimation des doses individuelles

Quand l'exposition mensuelle est inférieure au niveau de sensibilité du film ou au seuil d'enregistrement, la dose individuelle est alors considérée comme nulle. Une telle pratique est totalement justifiée pour les très nombreux travailleurs qui doivent être surveillés bien qu'ils ne soient jamais directement exposés. D'autre part, l'expérience montre que ne se trouvent jamais des relevés inférieurs au seuil chez les agents dont les travaux sont de nature répétitive (fabrication, contrôle) dans un environnement où le niveau d'exposition est clairement différent de zéro (de 5 à 10 mSv par an). Par contre, cette pratique conduit à une sous-estimation de la dose chez le travailleur effectivement exposé, mais de manière faible.

Pour de tels travailleurs, existe-t-il une démarche meilleure que celle de mettre à zéro tout relevé inférieur au seuil d'enregistrement ? Il est possible d'en douter lorsque les doses sont dues à un travail ponctuel dans le temps et dans l'espace, ce qui correspond assez classiquement au cas du chercheur qui met au point un procédé lors d'une campagne d'expériences. Faire mieux est possible si le travail effectué présente une certaine régularité et si chaque travailleur possède, en moyenne, au moins un relevé dosimétrique sur deux qui soit supérieur au seuil d'enregistrement. Dans cette situation, toute valeur nulle peut être alors considérée

comme une valeur manquante pour laquelle se pose un problème d'estimation. Voici quelques modes d'estimation classiques :

. Méthode A

Le relevé inférieur au seuil est mis à zéro ; déjà mentionnée ci-dessus, c'est la pratique usuelle ; elle entraîne une sous-estimation de la dose.

. Méthode B

Le relevé inférieur au seuil prend la valeur du seuil ; cette attitude pessimiste peut entraîner une surestimation de la dose.

. Méthode C

Le relevé inférieur au seuil prend une valeur égale à la moitié du seuil ; cette pratique est caractéristique de la situation d'ignorance totale ; en effet, les doses observées ne sont pas prises en compte et toutes les doses inférieures au seuil d'enregistrement sont considérées comme équiprobables.

2.2. Les pratiques fondées sur une analyse statistique

Eviter ces pratiques classiques implique de recourir à des méthodes plus sophistiquées qui sont évidemment plus lourdes à mettre en œuvre. Plusieurs solutions sont possibles à partir soit des distributions individuelles de doses mensuelles, soit des distributions mensuelles des doses individuelles du groupe étudié, soit de la distribution globale de toutes les doses mensuelles du groupe. Choisir une manière plutôt qu'une autre est, bien entendu, fonction des hypothèses adoptées et des informations possédées. Le relevé d'un individu, inférieur au seuil pour un mois donné, est estimé dans le premier mode à partir de ses propres doses mensuelles supérieures au seuil. Dans les deux autres modes, l'individu devient un membre d'un groupe étudié durant un mois particulier ou sur une période donnée. Mais il faudra, de toute façon, associer aux distributions observées des distributions théoriques, c'est-à-dire recourir à des lois de probabilité.

Dans le cas où la distribution des doses mensuelles d'un sujet est retenue afin de compléter une donnée qui lui manque, Jammet et al. [5] ont montré que la loi normale (Laplace - Gauss), dont la distribution est symétrique par rapport à la moyenne, constituait un modèle acceptable. Il faut souligner que, dans ce cas, les données sont toujours en petit nombre, au moins sur un an, au plus sur quelques années. Dès lors, deux procédures sont possibles :

. Méthode D1

Déterminer la valeur la plus probable de la moyenne d'une loi normale qui conduise au nombre de données manquantes observées ; dans la loi ainsi définie, la valeur attendue (en fait, l'espérance mathématique) de la dose est déterminée sachant qu'elle est inférieure au seuil d'enregistrement et la ou les données manquantes sont remplacées par cette valeur.

. Méthode D2

La même démarche peut être appliquée en utilisant, non plus le nombre de données manquantes, mais la moyenne des valeurs observées supérieures au seuil.

. Méthode E

Souvent, une autre méthode, notée ici méthode E, est utilisée. Elle procède de manière plus globale : le travailleur dans son groupe n'est qu'un individu parmi d'autres, l'avantage immédiat étant d'augmenter considérablement le nombre de doses mensuelles. Les disparités entre doses conduisent, en général, à une distribution observée qui est dissymétrique à gauche, c'est-à-dire dont la moyenne est supérieure à la valeur médiane. Le modèle classiquement utilisé est le modèle log-normal. Les problèmes d'estimation des paramètres de cette loi en présence de données manquantes sont abordés dans Gilbert et Kinnison [3]. L'estimation des doses manquantes nécessite un calcul de la dose attendue sachant qu'elle est inférieure au seuil ; par exemple, il est possible de la calculer en reportant d'abord toutes les doses mensuelles du groupe étudié en coordonnées gaussio-logarithmiques puis en extrapolant la courbe observée aux valeurs inférieures au seuil d'enregistrement.

Cette modélisation par la loi log-normale est jugée inadéquate par Darby et al. [2] qui lui préfèrent la loi de Weibull ; avec le même nombre de paramètres, deux, cette distribution présente, en effet, une plus grande variété de formes et offre aussi la possibilité d'effectuer une extrapolation graphique.

Les méthodes d'estimation notées A, B, D1, D2 et E ont été comparées par Jammet et al. [5] sur cinq petits groupes de travailleurs exposés durant un an. L'indicateur choisi pour comparer globalement les résultats est la dose collective annuelle en hommes-sieverts. Les tableaux I et II, extraits de [5], donnent respectivement des informations sur les cinq groupes étudiés et le résultat de la comparaison.

TABLEAU I
Effectif et dose mensuelle moyenne des cinq groupes étudiés, d'après [5].

Unité	Nombre de travailleurs	Dose mensuelle moyenne (mSv)	Lectures (%) inférieures au seuil (0,15 - 0,20 mSv)
Marcoule dégainage	83	1,03	7,3
Marcoule décontamination	147	1,46	11,2
La Hague décontamination	77	1,34	17,6
FAR ⁽¹⁾ - transuraniens	26	0,43	28,7
Cadarache Rapsodie	123	0,11	91,6

(1) FAR : Fontenay-aux-Roses.

Lois	Densité de probabilité	Moyenne	Variance
Lognormale	$\frac{1}{tb \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(t-a)}{b}\right)^2\right) \quad t > a \quad b > 0$	$\exp\left(a + \frac{b^2}{2}\right)$	$e^{2a+b^2}(e^{b^2}-1)$
De Weibull	$\frac{\beta}{\eta} \left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^\beta\right) \quad t > \gamma \quad \eta > 0 \quad \beta > 0$	$\eta I\left(1 + \frac{1}{\beta}\right) + \gamma$	$\eta^2 \left[I\left(\frac{2}{\beta} + 1\right) - I^2\left(1 + \frac{1}{\beta}\right) \right]$

Fig. 1. - Densités de probabilité des lois log-normale et Weibull.

TABLEAU II
Comparaison des méthodes d'estimation, d'après [5].

Unité	Dose collective annuelle (h.Sv)				
	A	B	D1	D2	E
Marcoule dégainage	0,940	0,948	0,949	0,948	0,952
Marcoule effluents décontamination	2,360	2,382	2,384	2,372	2,383
La Hague décontamination	1,136	1,147	1,146	1,139	1,142
FAR ⁽¹⁾ - transuraniens	0,122	0,137	0,129	0,133	0,134
Cadarache Rapsodie	0,147	0,364	?	?	?

(1) FAR : Fontenay-aux-Roses.

D'après le tableau II, les enseignements de cette application sont les suivants :

- si les expositions mensuelles inférieures au seuil d'enregistrement (0,25 mSv) sont peu nombreuses, les différentes méthodes conduisent à des résultats analogues ;
- si les doses non nulles sont en faible majorité et ne sont que de peu supérieures au seuil, des différences de l'ordre de 10% peuvent apparaître sur la dose collective annuelle ;
- aucune méthode ne peut s'appliquer dans les groupes où l'exposition des travailleurs est rare, relativement faible et mal répartie dans le temps ; la solution sage est alors fournie par la méthode A qui est la méthode usuelle.

Une autre question relative au seuil d'enregistrement a souvent été soulevée : jusqu'à quel niveau ce seuil peut-il être relevé sans réelle perte d'information dosimétrique, les doses inférieures au niveau choisi étant considérées comme nulles et non enregistrées ? Les niveaux possibles peuvent être pris de 0,3 mSv à 0,5 mSv (1/10 de la limite de la dose annuelle pour un individu du public) ou même dans des valeurs qui sont plus élevées. Il ne s'agit évidemment plus ici de compléter un fichier de relevés dosimétriques individuels mais plutôt de le réduire en n'enregistrant que les doses jugées élevées, le bénéfice attendu étant surtout de diminuer de façon substantielle le travail de nature administrative qu'implique la surveillance. La perte d'information dosimétrique est jugée à l'aide de la dose collective annuelle des groupes surveillés. Ilari et Johnston [4] ont abordé cette question en étudiant 18 groupes homogènes de travailleurs ; ils concluent que le seuil pourrait être augmenté de manière conséquente sans réelle perte d'information dosimétrique. L'intérêt peut apparaître pour un fichier des relevés tenu manuellement. En revanche, la gestion du fichier, surtout s'il est informatisé, se complique par le fait que seules les doses mensuelles supérieures au seuil y seraient incorporées et les agents n'auraient plus sur l'année le même nombre de relevés mensuels. Aussi le gain réel de gestion ne peut venir que du déclassement des agents non exposés.

3. LE FICHIER DES DOSES INDIVIDUELLES ANNUELLES

La donnée de base est la dose individuelle annuelle, cumul des 11 relevés dosimétriques mensuels. Dans ce cumul, les relevés mensuels inférieurs au seuil d'enregistrement sont actuellement considérés comme nuls. Pour une activité et une année données, la distribution des doses individuelles annuelles fournit aux gestionnaires d'installation le profil dosimétrique de l'activité. Classiquement, l'état de la radioprotection s'apprécie chaque année au moyen des deux indicateurs usuels qui sont la dose individuelle moyenne (celle-ci pouvant être calculée soit en considérant l'effectif total, soit l'effectif des individus à dose non nulle) et la dose collective pour le groupe concerné. Ces indicateurs ont un sens clair lorsque les deux conditions suivantes sont réunies :

- le groupe est homogène, le travail étant de même type pour tous ;
- les sujets sont des permanents, c'est-à-dire présents toute l'année dans l'installation.

Dans ce cas, la distribution des doses est plutôt symétrique et la moyenne est bien le paramètre adéquat pour résumer les données. Dans le cas contraire — hétérogénéité des doses suivant les postes de travail et personnel temporaire — d'autres indicateurs plus spécifiquement orientés vers la gestion de la radio-

protection sont utilisés :

- dose moyenne par an chez les sujets dont la dose annuelle n'est pas nulle (Cf. ci-dessus) ; c'est implicitement reconnaître que le groupe étudié est hétérogène car comprenant des travailleurs non exposés,
- dose moyenne par homme-an pour les activités où du personnel temporaire est requis,
- dose par mégawatt électrique,
- dose par tonne de minerai extrait,
- etc.

Quand il s'agit de comparer des activités, la démarche usuelle passe d'abord par la modélisation des distributions observées.

Dans un premier temps, le tracé de l'histogramme des doses suggère une famille de lois de probabilité, et le tracé de la fonction de répartition donne immédiatement la proportion des individus qui, dans le groupe, ont des expositions au-dessus d'une certaine valeur ou bien comprises dans un intervalle donné.

Dans un second temps, un test d'ajustement à une loi théorique est réalisé. La loi normale est parfaitement justifiée si les groupes sont homogènes. Compte tenu du fait que l'hétérogénéité constitue le cas courant, l'UNSCEAR [7] a proposé l'ajustement à une loi log-normale ; mais l'expérience montre que ce modèle est souvent inadéquat et Kendall et al. [6] par exemple, en préfèrent d'autres.

La variété des situations est telle qu'*a priori* de nombreux modèles sont possibles et donc toute tentative pour imposer un modèle unique est vouée à l'échec. Pour l'objet de la comparaison, il semble donc suffisant de reporter les fonctions de répartition des groupes étudiés sur un seul graphique. L'exigence de lisibilité guide le choix des échelles : l'ordonnée — en échelle arithmétique ou en échelle gaussienne — est associée au report direct des probabilités cumulées, l'abscisse — arithmétique ou logarithmique — sert au report des doses ou de leurs logarithmes.

4. DOSE-VIE

4.1. Position du problème

Pour le travailleur, la dose-vie n'est que le cumul de ses doses annuelles durant sa vie professionnelle. Par contre, pour les besoins de la radioprotection, la dose-vie du travailleur exerçant son activité au sein d'un groupe donné (défini par un poste de travail, un atelier ou plus largement une installation) pose un problème d'estimation. Pour une vie de travail, ou plus simplement pour une période de travail au sein du même groupe de travailleurs, la pratique usuelle est de sommer les moyennes annuelles qui résument les doses individuelles annuelles du groupe considéré pour la période étudiée. Cette pratique est tout à fait justifiée si le groupe est homogène quant aux postes de travail et quant à l'âge des individus le composant. C'est dans un tel contexte que la possibilité de fixer une limitation sur la dose-vie du travailleur et non sur la dose annuelle, peut être évoquée. Ce n'est pas, cependant, le choix retenu par la CIPR et dans les diverses réglementations nationales et européennes, à la fois pour des raisons de commodité et de protection. En effet, il est techniquement plus simple d'assurer

une protection individuelle en ne surveillant qu'une dose annuelle bien que cela se révèle plus contraignant pour l'exploitant dans d'assez nombreuses situations ; donc, le pari de faire en sorte que la dose cumulée sur une période longue n'ex-cède pas le cumul correspondant des limites annuelles autorisées est évité, pari qui peut être difficile à tenir si la carrière de l'agent présente une nette surexposition dans les premières années.

Quoi qu'il en soit, le problème d'estimer la dose individuelle reçue en moyenne sur une période fixée reste à résoudre, dès lors que se manifestent les phénomènes de spécialisation dans un poste, de vieillissement des individus, de glissement d'un poste de travail à un autre. Il faut alors suivre une cohorte de travailleurs, c'est-à-dire les mêmes sujets tout au long de la période considérée. Chaque travailleur apporte son suivi dosimétrique et, *a priori*, plus long est son suivi, meilleure sera l'estimation. En fait, un compromis doit être réalisé entre la taille de la cohorte (nombre de sujets) et la durée de la période étudiée : en effet, plus longue est la période, plus petite sera la taille de la cohorte (perte inévitable des sujets pour des motifs divers), et donc la qualité des études statistiques faites en pâtit, les estimations de la dose cumulée devenant moins précises.

4.2. Traitement statistique sur un exemple

Quatre groupes suivis au plan de l'exposition externe gamma ont été choisis pour saisir de manière statistique les phénomènes mentionnés ci-dessus :

Groupe MU : Mine d'uranium ; période 1971-1979 ; cohorte de 449 mineurs de fond dans une population d'effectif annuel moyen de 1000, avec un effectif minimal de 631 et maximal de 1484.

Groupe RP : Réacteur de puissance à eau lourde ; postes autres que les services généraux ; période 1971-1981 ; cohorte de 78 travailleurs dans une population d'effectif annuel moyen de 200 avec un effectif minimal de 109 et maximal de 540.

Groupe EC : Fabrication d'éléments combustibles au plutonium ; période 1971-1978 ; cohorte de 72 travailleurs dans une population d'effectif annuel moyen de 150 avec un effectif minimal de 110 et maximal de 202.

Groupe TU : Fabrication de sources de transuraniens à usage industriel ; période 1974-1981 ; cohorte de 25 travailleurs dans un atelier dont l'effectif annuel moyen est de 55 agents avec un minimum de 50 et un maximum de 64 agents.

En notant :

D_j^i : la dose annuelle du sujet i pour l'année j

n : le nombre de sujets

y : le nombre d'années de suivi de la cohorte

alors les indicateurs statistiques sont les suivants :

— La moyenne des doses annuelles de l'individu i dans la période considérée qui s'écrit :

$$\overline{D^i} = \frac{1}{y} \sum_{j=1}^y D_j^i$$

Cet indicateur est équivalent à la dose cumulée du sujet pour la période. Son avantage est de pouvoir s'abstraire de la durée de suivi; il tient compte des spécificités du sujet.

— La dose individuelle moyenne de l'année j est l'indicateur traditionnellement utilisé pour juger de l'évolution de la radioprotection :

$$\bar{D}_j = \frac{1}{n} \sum_{i=1}^n D_j^i$$

La dispersion des doses individuelles de l'année j est classiquement mesurée au moyen de l'écart-type σ_j

$$\sigma_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (D_j^i - \bar{D}_j)^2}$$

Bien que les distributions des doses individuelles annuelles ne soient généralement pas symétriques, la comparaison des écarts-types annuels permet de juger de la plus ou moins grande dispersion des doses au cours des années.

— La liaison entre les doses annuelles successives est mesurée par les coefficients de corrélation linéaire entre les doses reçues par le groupe au cours d'une année et celles qu'il reçoit au cours des années ultérieures: pour les années j et k , le coefficient de corrélation s'écrit :

$$r_{j,k} = \frac{\frac{1}{n} \sum_{i=1}^n (D_j^i - \bar{D}_j)(D_k^i - \bar{D}_k)}{\sigma_j \cdot \sigma_k}$$

$r_{j,k}$ est toujours compris entre -1 et $+1$.

Il est intéressant de voir ce qui se passe d'une année sur l'autre ou avec un délai de deux ans, ou plus,... Ceci conduit à prendre les estimations suivantes pour la corrélation linéaire entre les doses reçues :

sur deux années consécutives :

$$r_2 = \frac{1}{y-1} \sum_{j=1}^{y-1} r_{j,j+1}$$

et sur k années d'intervalle :

$$r_k = \frac{1}{y-k} \sum_{j=1}^{y-k} r_{j,j+k}$$

Quel est le sens de ces indicateurs de liaison ?

Si les doses reçues annuellement le sont indépendamment d'une année sur l'autre, alors tous les coefficients de corrélation linéaire doivent être proches de zéro. Par contre, plus les valeurs de r tendent vers 1, plus les doses sont reçues de manière analogue par les sujets composant la cohorte; c'est-à-dire que, sur la période étudiée, les sujets le plus souvent exposés le restent, les moins exposés le sont encore de même que les moyennement exposés. Pour des valeurs de r qui tendraient vers -1 , cas très improbable, un ordre inverse s'établirait où le

travailleur avec la dose la plus élevée deviendrait le moins exposé, etc.

Ces indicateurs constituent donc un moyen simple pour apprécier si les doses sont distribuées aléatoirement au cours du temps.

Le tableau III récapitule les doses individuelles moyennes dans chacun des groupes et dans la cohorte qui en est extraite pour les années de suivi. L'évolution des moyennes de l'exposition externe gamma traduit grossièrement le niveau de radioprotection. On note que les évolutions dans le groupe et dans la cohorte associée sont similaires; ceci est normal car chaque cohorte est de taille conséquente relativement aux effectifs annuels du groupe dont elle est issue. Dans le groupe RP, EC, TU, la dose individuelle moyenne dans la cohorte est en léger excès par rapport à celle du groupe pris dans son ensemble. La situation ne se reproduit pas chez les mineurs et il serait même possible de trouver d'autres groupes où les membres de la cohorte, qui peuvent être considérés comme des permanents, seraient, en moyenne, nettement moins exposés que l'individu moyen du groupe pour chaque année.

Un tel tableau n'apporte aucune information sur la distribution des doses annuelles et sur une éventuelle spécialisation dans le travail. Pour ce faire, il faut utiliser les fonctions de répartition (ou courbe de fréquences cumulées) des distributions annuelles. La fonction de répartition donne immédiatement la proportion des travailleurs dont les expositions dépassent une valeur donnée ou appartiennent à un intervalle de dose pré-défini. Dans la figure 2, sont représentées trois fonctions de répartition construites à partir des seules données de la cohorte des 449 mineurs de fond issus du groupe MU.

Fig. 2. - Fonctions de répartition des doses construites pour les trois situations suivantes :

- ① toutes les doses de 1971 à 1979;
- ② moyennes individuelles des doses reçues de 1971 à 1979;
- ③ moyennes individuelles des doses si elles étaient reçues de façon aléatoire de 1971 à 1979. Cohorte: 449 sujets; $X + \gamma$: moyenne 6,27 mSv.

TABLEAU III

Doses individuelles moyennes, en mSv, dans chacun des groupes et dans la cohorte

Année	MU				RP			EC			TU		
	Effectif total	$\bar{D}$	$\bar{D}$ Cohorte 449 sujets	σ	Effectif total	$\bar{D}$	$\bar{D}$ Cohorte 78 sujets	Effectif total	$\bar{D}$	$\bar{D}$ Cohorte 72 sujets	Effectif total	$\bar{D}$	$\bar{D}$ Cohorte 25 sujets
1971	696	8.8	8.7	6.8	269	.4	.9	157	2.6	3.3			
1972	631	9.7	9.1	7.8	164	2.2	3.	128	4.1	5.			
1973	640	7.8	7.2	6.65	169	2.6	3.2	110	2.6	3.			
1974	673	6.7	6.2	6.	138	2.2	3.2	110	2.6	4.2	51	15.6	18.6
1975	960	4.9	4.9	5.5	129	2.3	3.	126	2.9	3.7	51	13.4	16.5
1976	1094	5.1	5.4	5.5	132	2.7	3.	114	2.3	2.6	53	13.4	16.6
1977	1250	4.5	5.2	5.	156	4.9	4.7	111	2.	2.	54	15.8	16.8
1978	1372	4.5	5.3	5.8	109	4.3	4.9	162	1.7	2.	58	12.	13.9
1979	1484	4.1	4.4	4.9	153	2.5	3.1				64	8.	9.9
1980					201	3.8	4.2				63	6.8	7.8
1981					216	5.	5.5				63	7.7	9.

σ : Ecart-type des doses individuelles annuelles.

La fonction de répartition ① a été obtenue en regroupant les 4041 (= 449 x 9) doses reçues dans la cohorte au cours des 9 années. Implicitement, l'unité statistique utilisée est l'homme-an et cette procédure d'agrégation est d'autant plus justifiée que d'une année sur l'autre les distributions de doses ne changent guère; en effet, dans le cas contraire, la distribution résultant du mélange des doses est, à coup sûr, extrêmement dispersée et ne peut servir à modéliser une dose-vie.

La fonction de répartition ② est celle des 449 moyennes individuelles ($\bar{D}_i$, $i = 1, 449$). Les deux distributions mentionnées ont même moyenne, 6,27 mSv, mais les maximums diffèrent: pour ①, 42,8 mSv est la plus forte dose annuelle reçue pendant ces 9 années et pour ②, 23 mSv correspond à l'individu le plus exposé en moyenne sur les 9 ans, sa dose cumulée de 1971 à 1979 étant de $9 \times 23 \text{ mSv} = 207 \text{ mSv}$. Il est normal que les moyennes individuelles prises sur une assez longue période soient plus concentrées que les doses annuelles et donc plus aisément au-dessous d'une valeur limite que ne pourraient l'être des doses annuelles: ainsi, dans la figure 2, 83 doses annuelles dépassent 23 mSv tout en restant inférieures à 42,8 mSv. Quant à la fonction de répartition ③, elle correspond à la situation hypothétique suivante: les 449 mineurs de fond de la cohorte constitueraient un groupe homogène relativement aux tâches réalisées et, donc, quant aux doses reçues. Celles-ci seraient alors distribuées chaque année au hasard parmi les travailleurs et aussi d'une année sur l'autre; la moyenne individuelle des 9 doses reçues se comporte en conséquence comme la moyenne de 9 variables stochastiquement indépendantes dont la distribution est modélisée de façon satisfaisante par une loi normale de moyenne 6,27 mSv et d'écart-type 2,02 mSv,

$$2,02 \text{ mSv} = \sqrt{\frac{1}{9} \sum_{j=1}^9 \sigma_j^2}$$

où les σ_j sont ceux donnés dans le tableau III. Comme le montre la figure 2, l'écart entre cette situation hypothétique illustrée par la courbe ③ et la situation réelle donnée par la courbe ② est très important. Dans l'hypothèse d'indépendance, il ne devrait y avoir que 7 sujets environ sur 449 (soit 1,5%) dont la dose dépasserait 10 mSv par an et ce, sur les 9 années, alors que dans la réalité ils sont 90 (soit 20%) à être dans ce cas là. Ceci est dû à la forte liaison entre les doses successives reçues par chaque travailleur. Ce dernier constat existe pour d'autres catégories présentées par Brenot et al. [1]. Ici, il est mis en évidence dans le tableau IV qui donne les estimations des coefficients de corrélation linéaire entre la dose D_i reçue pendant l'année j , et la dose D_{j+1} et D_{j+2} , etc. Ces liaisons temporelles sont donc relatives au nombre d'années d'écart entre les diverses doses.

Ainsi, pour les cohortes MU, RP, TU et, à un degré moindre, EC, une forte corrélation linéaire s'explique essentiellement par la spécialisation dans le travail bien que les groupes aient été choisis de façon à être aussi homogènes que possible. Les cohortes étudiées et donc les groupes dont elles sont issues présentent, au sein d'une même activité, des tâches encore très diversifiées. Ces liaisons ont pour conséquence que la distribution des doses individuelles reçues en moyenne par an sur une période assez longue ne peut être correctement modélisée à partir de l'approche suivante:

- dans un premier temps, calculer la moyenne et l'écart-type de toutes les doses annuelles de la période;
- puis retenir un modèle gaussien dont les paramètres sont la moyenne calculée et l'écart-type divisé par la racine carrée du nombre d'années observées.

En effet, cette démarche suppose implicitement que les doses soient, chaque année, distribuées aléatoirement au sein du groupe de travailleurs, ce qui est bien rarement le cas.

TABLEAU IV
Estimation des coefficients de corrélation linéaire

Ecart (ans)	1	2	3	4	5	6	7	8	9	10
Cohorte MU	0,78	0,68	0,63	0,59	0,56	0,56	0,57	0,52		
Cohorte RP	0,77	0,71	0,70	0,71	0,72	0,65	0,68	0,73	0,69	0,61
Cohorte EC	0,75	0,61	0,46	0,30	0,23	0,19	0,13			
Cohorte TU	0,73	0,64	0,62	0,55	0,58	0,55	0,46			

4.3. Bilan

Ainsi, une prévision de dose cumulée sur une longue période est chose délicate à bien faire. A partir d'un suivi existant, la prévision pour un individu donné doit s'appuyer sur sa propre histoire dosimétrique; celle pour l'individu anonyme au sein d'un groupe requiert d'employer une cohorte issue du groupe pour tenir compte à la fois du vieillissement des agents et d'éventuelles modifications apportées à la surveillance; mais, plus le groupe est homogène et moins le choix d'une cohorte s'impose. Dans le cas de nouveaux ateliers ou d'installations sans passé dosimétrique, la prévision souvent faite à partir d'un scénario pessimiste conduit à la mise en place de protections excessives et donc à des surcoûts qui ne sont pas toujours justifiés.

5. CONCLUSION

La surveillance dosimétrique des travailleurs, lorsqu'elle dépasse le cadre réglementaire de la simple confrontation des résultats dosimétriques mensuels et des cumuls annuels aux limites autorisées, pose aux responsables de radioprotection des problèmes d'estimation qui ne sont pas pris en compte dans les procédures standardisées. En effet, s'il faut partir des doses individuelles pour estimer une dose manquante, ou, au niveau des cumuls annuels, établir le profil dosimétrique d'un groupe, suivre son évolution et le comparer à d'autres, plusieurs possibilités se présentent. Mais, pour toutes ces estimations, il est

nécessaire de rechercher la forme de la distribution statistique associée aux données. Si cette distribution est symétrique, sa modélisation par la loi normale est justifiée et les extrapolations sont aisées ; de plus, les indicateurs usuels ont alors une signification claire. Par contre, si la distribution est asymétrique, le recours à une autre loi de probabilité (loi log-normale, loi de Weibull) implique un choix qui n'est pas évident et les indicateurs habituels perdent en grande partie leur validité.

Quant au problème de l'estimation de la dose individuelle moyenne sur une longue période, il exige, pour être résolu, qu'une cohorte soit extraite du groupe étudié et qu'elle soit suivie suffisamment longtemps. Les indicateurs de corrélation utilisés conjointement avec les fonctions de répartition permettent alors d'apprécier correctement l'évolution de la situation dosimétrique.

Il résulte des exemples étudiés que les pratiques habituelles basées sur l'emploi d'indicateurs statistiques simples ont une capacité de description limitée aux situations classiques, et il convient donc de ne pas l'oublier avant de les utiliser pour l'interprétation des données. Tout approfondissement de l'analyse dans une perspective de gestion du risque radiologique ne peut se faire qu'au prix d'une réflexion sur la situation considérée, qui débouche, en général, sur une solution satisfaisante et relativement simple, sans qu'il soit obligatoirement nécessaire de recourir à des techniques statistiques complexes.

6. RÉFÉRENCES

- [1] BRENOT J., PIECHOWSKI J., TIRMARCHE M. Statistical analysis of results on miners' exposures. In: Radiation Hazards in mining, Golden, Oct.4-9, 1981. (GOMEZ M., Ed.), New York: Society of mining engineers, 1981, 731-737.
- [2] DARBY S.C., KENDALL G.M., GREENSLADE E. Patterns of dose incurred by workers on the NRPB dose record keeping service, J. Soc. Rad. Prot., 1982, 2 (4), 31-38.
- [3] GILBERT R.O., KINNISON R.R. Statistical methods for examining the mean and variance from radionuclide data sets containing negative unreported or less-than values. Health Phys., 1981, 40, 377-390.
- [4] ILARI O., JOHNSTON P. Study on the distribution of individual doses to workers. ICRP Committee 4 Internal report, 1984.
- [5] JAMMET H., MECHALI D., NENOT J.C., BRENOT J. Distribution des doses individuelles et interprétation statistique. In: The dose limitation system in the nuclear fuel cycle and in radiation protection, Madrid, 19-23 Oct. 1981. Vienne: AIEA, 1982, 363-375.
- [6] KENDALL G.M., DARBY S.C., GREENSLADE E. Patterns of dose incurred by workers on the NRPB dose record keeping service - annual doses. J. Soc. Rad. Prot., 1982, 2 (3), 20-25.
- [7] UNITED NATIONS SCIENTIFIC COMMITTEE ON THE EFFECTS OF ATOMIC RADIATION (UNSCEAR). Sources and effects of ionizing radiation. New York: United Nations, 1977.