

Chapitre 14

Robustesse d'une typologie des pays développés suivant leur profil de décès par cause : 1968-1984

Nicolas BROUARD
Institut national d'études démographiques (INED)
Paris, France

Le développement de la théorie statistique a son origine dans l'étude de la mortalité. La mesure des risques de décès a tout d'abord intéressé les assureurs ou les banquiers. En effet, une estimation correcte de la mortalité par âge permet de parier sur la vie des hommes par le principe de l'assurance, et le service rendu par cette assurance est une source de profit. La littérature abonde donc dans le domaine de l'estimation, l'ajustement et la modélisation de la mortalité par âge. Mais, pour un actuairé, il est moins important de connaître de quoi va décéder un client que sa date de décès. Aussi, la connaissance de la cause de décès intéresse un autre secteur, celui des organismes de santé. Force est alors de constater que, dans le domaine de la modélisation de la mortalité par cause de décès, la littérature est nettement plus pauvre. On doit invoquer aussi la moindre abondance et la moindre fiabilité des données, mais il est vrai que le sujet est difficile et déroutant. Déjà, nous avons appris avec les tables types de Ledermann⁽¹⁾ et de Coale et Demeny⁽²⁾ que deux pays peuvent avoir la même espérance de vie, mais des profils de risques de décès par âge tout à fait différents. Avec la mortalité par cause, l'analyse se complique beaucoup plus, car deux pays qui ont des taux de mortalité de niveau identique à certains âges comme le Japon ou la Suède, ont des profils par cause radicalement opposés : d'une manière succincte, on meurt de cancer de l'estomac et de maladies cérébro-vasculaires au Japon, et de maladies cardio-vasculaires en Suède.

Dans un article présenté avec Alan Lopez au Congrès de Florence⁽³⁾, nous avons essayé de définir une typologie des pays à basse mortalité selon la cause de

(1) LEDERMANN (Sully) et BRÉAS (Jean). — Les dimensions de la mortalité. — *Population*, 1959, n° 4.

(2) COALE (Ansley J.) et DEMENY (Paul). — *Regional Model Life Tables and Stable Populations*. — Princeton, N.J., Princeton University Press, 1966.

(3) BROUARD (Nicolas) et LOPEZ (Alan). — Causes of deaths patterns in low mortality countries : a classification analysis. In : *Congrès international de la population, Florence, 5-12 June 1985*. — Liège, UIESP, Ordina Éditions, 1985, p. 385-406.

Chapitre 14

Robustesse d'une typologie des pays développés suivant leur profil de décès par cause : 1968-1984

Nicolas BROUARD
Institut national d'études démographiques (INED)
Paris, France

Le développement de la théorie statistique a son origine dans l'étude de la mortalité. La mesure des risques de décès a tout d'abord intéressé les assureurs ou les banquiers. En effet, une estimation correcte de la mortalité par âge permet de parier sur la vie des hommes par le principe de l'assurance, et le service rendu par cette assurance est une source de profit. La littérature abonde donc dans le domaine de l'estimation, l'ajustement et la modélisation de la mortalité par âge. Mais, pour un actuair, il est moins important de connaître de quoi va décéder un client que sa date de décès. Aussi, la connaissance de la cause de décès intéresse un autre secteur, celui des organismes de santé. Force est alors de constater que, dans le domaine de la modélisation de la mortalité par cause de décès, la littérature est nettement plus pauvre. On doit invoquer aussi la moindre abondance et la moindre fiabilité des données, mais il est vrai que le sujet est difficile et déroutant. Déjà, nous avons appris avec les tables types de Ledermann⁽¹⁾ et de Coale et Demeny⁽²⁾ que deux pays peuvent avoir la même espérance de vie, mais des profils de risques de décès par âge tout à fait différents. Avec la mortalité par cause, l'analyse se complique beaucoup plus, car deux pays qui ont des taux de mortalité de niveau identique à certains âges comme le Japon ou la Suède, ont des profils par cause radicalement opposés : d'une manière succincte, on meurt de cancer de l'estomac et de maladies cérébro-vasculaires au Japon, et de maladies cardio-vasculaires en Suède.

Dans un article présenté avec Alan Lopez au Congrès de Florence⁽³⁾, nous avons essayé de définir une typologie des pays à basse mortalité selon la cause de

(1) LEDERMANN (Sully) et BRÉAS (Jean). — Les dimensions de la mortalité. — *Population*, 1959, n° 4.

(2) COALE (Ansley J.) et DEMENY (Paul). — *Regional Model Life Tables and Stable Populations*. — Princeton, N.J., Princeton University Press, 1966.

(3) BROUARD (Nicolas) et LOPEZ (Alan). — Causes of deaths patterns in low mortality countries : a classification analysis. In : *Congrès international de la population, Florence, 5-12 June 1985*. — Liège, UIESP, Ordina Éditions, 1985, p. 385-406.

décès. Cette typologie avait été établie à partir d'une analyse transversale de 45 pays dont l'année moyenne était l'année 1980. Nous avons ainsi montré que la Suède est caractérisée par une prépondérance de maladies cardio-vasculaires. C'est assez paradoxal puisque le taux de mortalité par maladies cardio-vasculaires en Suède est actuellement un des taux les plus faibles du monde. En fait, la Suède a une espérance de vie des plus élevées au monde qui la place au second rang mondial, et c'est la raison pour laquelle la mortalité par maladies cardio-vasculaires y est une des plus faibles, mais, et c'est là notre hypothèse, lorsque la masse des pays atteindra le même niveau d'espérance de vie, la Suède apparaîtra comme un pays où la mortalité par maladies cardio-vasculaires est une des plus fortes.

Nous tentons ici de confirmer la validité de cette typologie en testant sa robustesse à l'évolution de la mortalité à moyen terme. L'étude chronologique porte sur l'ensemble des données de mortalité disponibles à l'OMS concernant les pays développés dont les décès par cause sont établis suivant les 8ème et 9ème révisions de la Classification internationale des maladies. La première année est généralement l'année 1968, et la dernière, 1984, l'étude couvre ainsi une période de 15 ans. On cherchera aussi à évaluer la robustesse de la méthode en analysant les variations géographiques, à l'intérieur d'un même pays.

Avant de présenter les résultats, nous rappellerons brièvement la méthodologie et surtout les hypothèses.

I. — LES DONNÉES, LE MODÈLE ET SES HYPOTHÈSES

Rappelons en premier lieu que seule la cause initiale⁽⁴⁾ de décès est répertoriée dans la statistique de l'OMS : le processus, les causes multiples, les circonstances qui conduisent à la mort nous sont donc inconnus, et limitent certainement l'intérêt de notre étude.

Mis à part l'URSS, tous les pays développés envoient leurs statistiques à l'OMS qui les publie : notre étude porte ainsi sur 35 pays développés. Si les statistiques sont disponibles suivant une liste⁽⁵⁾ en 150 postes, nous nous sommes limité à 16 causes particulières (voir le tableau A en annexe 1) et cela pour au moins deux raisons.

1) La première provient d'un souci de simplicité ; l'intercorrélation entre les causes est suffisamment complexe pour qu'on réduise le nombre des paramètres.

2) La seconde modère la première en ce sens qu'il nous apparaît fondamental de ne pas agréger des causes de maladies dont la fréquence d'apparition est très variable suivant les pays. Ainsi, les tumeurs sont éclatées en 4 rubriques : poumon, voies aéro-digestives supérieures, sein (pour les femmes) et prostate (pour les hommes), estomac. De plus toutes les tumeurs n'ont pas été retenues ; nous n'avons sélectionné que les cancers importants en nombre ou qui faisaient l'objet d'études actuelles de la part des organismes de santé, des médecins ou épidémiologistes.

(4) Généralement appelée *cause principale* en France.

(5) Il s'agit de la « liste A » pour la 8ème révision et de la « liste pour la mise en tableaux » pour la 9ème.

II. — LA MÉTHODE STATISTIQUE

Nous ne nous intéressons pas au *niveau* de la mortalité pour telle ou telle cause, mais aux proportions de décès par cause. C'est l'optique des tables types. Nous voulons en effet montrer qu'un pays a un profil par cause de décès, qui, sans être immuable dans le temps, reste très caractéristique, si bien que l'écart entre deux profils d'un même pays à deux dates différentes (1968-1984) est plus faible que celui de deux pays différents à une même date.

De plus, comme il est trop compliqué d'introduire simultanément l'évolution de ces profils par cause en fonction de l'âge⁽⁶⁾, nous avons choisi trois groupes d'âge : 15-24, 45-54 et 65-74 ans. La mortalité à 20 ans ne fait pas intervenir les mêmes causes que la mortalité aux âges élevés, et est traitée à part. Les causes de décès à 20 ans sont ainsi réduites à quatre :

- 1) les accidents de véhicule à moteur,
- 2) les suicides,
- 3) les autres accidents,
- 4) les autres causes;

leur étude statistique est donc plus aisée. A 50 et 70 ans, au contraire, les 16 causes jouent un rôle important. Enfin, le niveau moyen, pour l'ensemble des 35 pays développés, du taux de mortalité à 50 ans est 6 à 7 fois plus faible que celui à 70 ans et sa répartition par cause est très différente comme l'indique la figure 1. Les maladies cardio-vasculaires prennent alors une part beaucoup plus importante et les accidents une place moindre. La répartition par sexe est elle aussi très différenciée.

Nous pouvons voir sur ces distributions calculées comme moyenne de tous les pays au début (1968-1969), en fin de la période (1983-1984), que si le niveau moyen de la mortalité a baissé (de 475 à 432 pour cent mille pour les hommes de 70 ans par exemple), la structure n'a pas énormément évolué. Évidemment, ces distributions moyennes cachent de très grandes variabilités entre pays, mais nous avons montré que si un pays avait, pour un sexe donné, une particularité à un âge donné, il l'avait aussi à un autre âge, et pour l'autre sexe. Ces particularités n'ont pas forcément la même ampleur, mais sont suffisantes pour que les typologies ou regroupements de pays, observés à 50 ans et à 70 ans, pour les hommes et pour les femmes, puissent être considérés comme identiques.

Décrivons maintenant la méthode statistique que nous avons appliquée afin de comprendre comment les nouveaux résultats de l'analyse longitudinale vont s'articuler avec la coupe transversale. On trouvera en annexe 2 plus de détail sur les calculs ainsi que des résultats numériques complémentaires. Une formule et les paramètres correspondants permettent de localiser à volonté sur nos graphiques n'importe quel pays pour lequel on disposerait d'une statistique de causes de décès comparable.

(6) DUCHÈNE (Josiane). — *Un essai de modélisation de la répartition des décès selon l'âge et la cause dans les pays industrialisés*. — Louvain-La-Neuve, Université Catholique de Louvain, Département de Démographie, Cabay, 1980, vol. VI, 82 p.

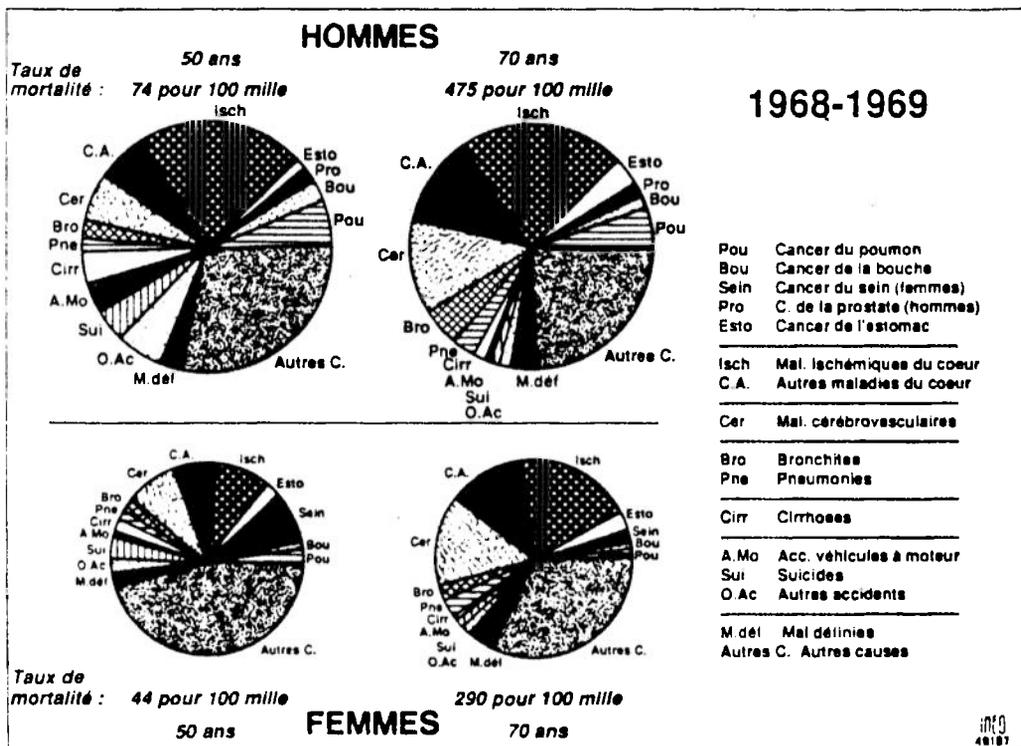
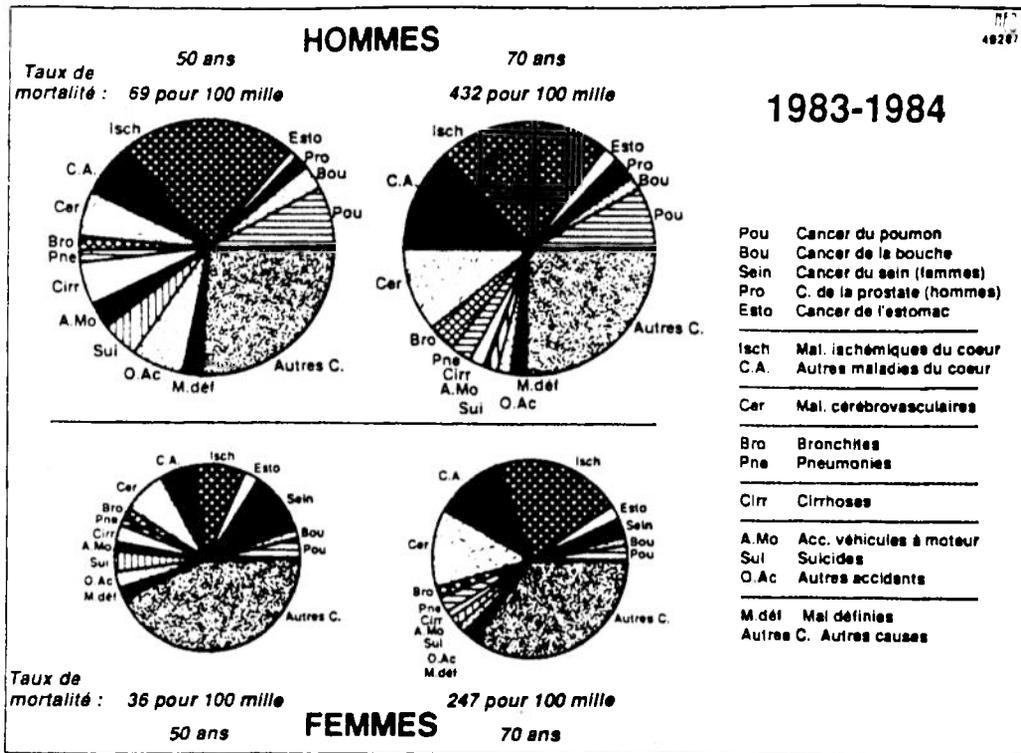


Figure 1. — Répartition des décès par cause à 45-54 ans et 65-74 ans pour les hommes et les femmes. Moyenne sur 35 pays en 1968-1969 et en 1983-1984

Considérons un tableau où sont portés en ligne les pays et en colonne les 16 causes choisies. Chaque case du tableau représente par exemple le taux de mortalité selon la cause j du pays i , pour un sexe et un âge donnés. La somme des taux en ligne donne alors le taux de mortalité générale. En divisant chacune des cases par ce taux de mortalité générale, nous obtenons un profil de décès par cause. Définissons alors une distance entre deux profils; calculons le tableau des distances entre pays 2 à 2, effectuons une classification automatique sur le tableau de distance, et nous obtenons une typologie. Evidemment, 1) il n'y a pas unicité de la distance, 2) les algorithmes de classifications abondent, 3) le nombre des groupes désiré est arbitraire. Nous avons néanmoins pu mesurer la robustesse de la typologie en changeant le type de distance et d'algorithme de classifications. La proximité entre groupes varie suivant les méthodes, mais la typologie proposée, sans être optimale pour chacun des cas, paraît tout à fait acceptable. Au lieu de faire une synthèse des différentes typologies obtenues, nous avons préféré donner précisément la méthode utilisée afin qu'elle puisse être reprise plus aisément.

Rappelons maintenant qu'au lieu de présenter deux analyses, l'une à 50 ans et l'autre à 70, nous avons (après avoir vérifié que les typologies concordent) juxtaposé les deux tableaux (voir l'annexe 2). La somme en ligne donne alors la somme des deux taux à 50 ans et à 70 ans. Pour donner de l'importance aux causes de moindre portée et rétablir l'équilibre entre les taux à 50 ans et à 70 ans, nous avons choisi la distance du χ^2 , qui est la somme des carrés des écarts entre les profils de deux pays, pondérée par l'inverse du poids *moyen* (sur tous les pays) de la cause à cet âge.

Nous aurions pu juxtaposer aussi les tableaux féminins et faire une typologie globale, mais nous avons préféré faire une analyse par sexe. La typologie est sensiblement identique, mais la représentation graphique de la typologie et sa correspondance avec les causes sont très différentes (en particulier les cancers du sein ne sont évidemment pas distribués d'une manière identique aux cancers de la prostate).

Si nous choisissons en effet la distance du χ^2 , l'analyse factorielle qui lui est associée offre une possibilité de *correspondance*⁽⁷⁾ entre les causes et les pays qui est très agréable et qui est la base de notre analyse longitudinale.

La typologie obtenue contenait 11 groupes. Les pays en voie de développement qui y figuraient ont été supprimés de l'analyse ultérieure, puisque notre sujet ne concerne que les pays développés. En voici la liste et la représentation sous forme de bulles sur les deux plans factoriels croisant respectivement les axes 1 et 2 (figure 2) et les axes 1 et 3 (figure 3) :

- groupe A : Japon;
- groupe B : Singapour, Hong-Kong;
- groupe C : France, (Porto Rico)⁽⁸⁾;

(7) BENZECRI (J.-P.) — *L'analyse des données, vol. 2 : l'analyse des correspondances*. 2ème édition. — Dunod, Paris, 1973 (vol. 2, seconde édition 1973).

(8) Les pays entre parenthèses sont les pays à basse mortalité mais non développés qui avaient été inclus dans la première étude (BROUARD (Nicolas) et LOPEZ (Alan). — *Causes of death patterns...*, op. cit., note 3).

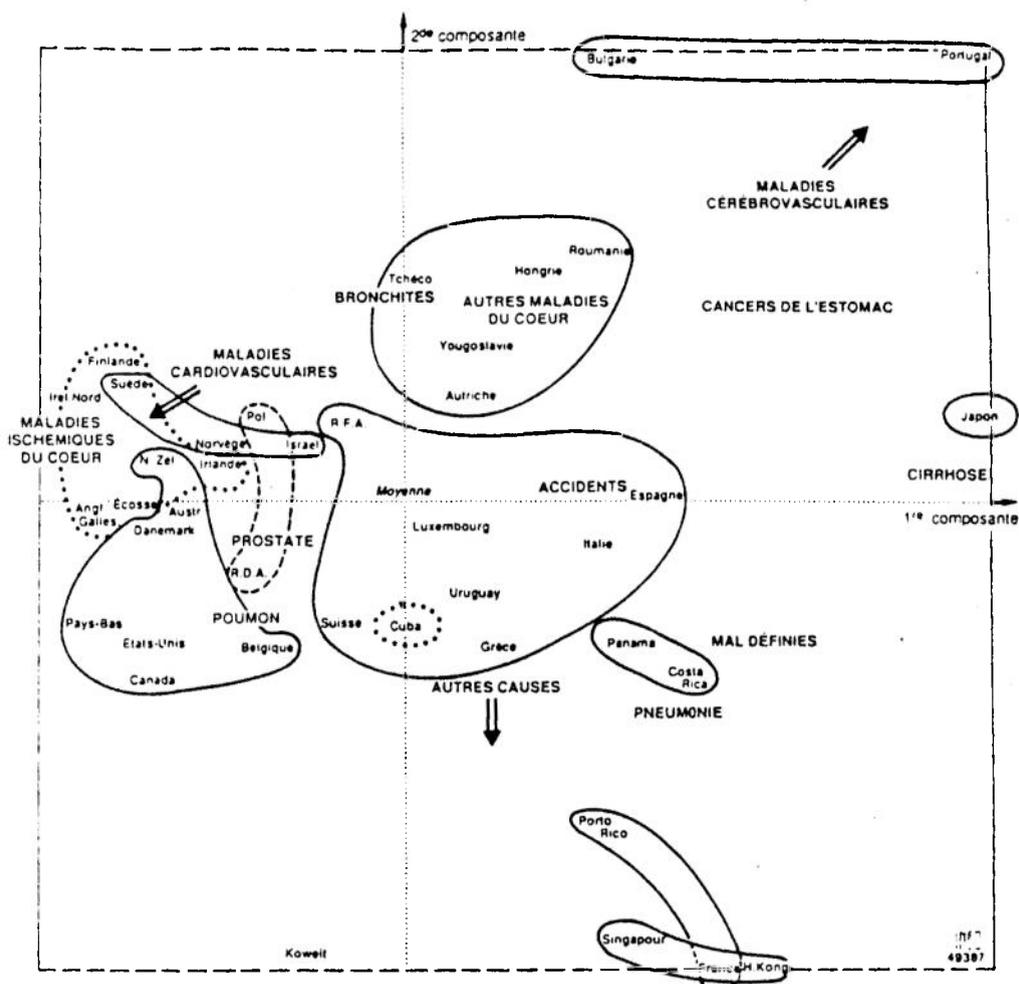


Figure 2. — Classification des pays selon leur profil de décès par cause à 45-54 ans et à 65-74 ans. Hommes. Première et seconde composantes

- groupe D : Portugal, Bulgarie;
- groupe E : (Panama, Costa Rica);
- groupe F : Espagne, Italie, Grèce, Uruguay, République fédérale d'Allemagne, Suisse;
- groupe G : Roumanie, Yougoslavie, Hongrie, Autriche, Tchécoslovaquie;
- groupe H : Pologne, République démocratique allemande;
- groupe I : États-Unis, Canada, Australie, Nouvelle-Zélande, Pays-Bas, Danemark, Belgique;
- groupe J : Angleterre et Galles, Écosse, Irlande, Irlande du Nord, Finlande, (Cuba);
- groupe K : Suède, Norvège, Israël.

Cette analyse factorielle oppose donc d'une part les pays à forte proportion de maladies cérébro-vasculaires, (coin Nord-Est de la figure 2) à ceux (Ouest) ayant une forte proportion de maladies cardio-vasculaires. L'axe 3 de la figure 3

discrimine alors, en haut, les décès dus à l'alcoolisme (cancer de la bouche et cirrhose) et en bas les pneumonies (Hong-Kong, Singapour).

Chaque nouvel axe discrimine d'autres maladies qui sont associées à des pays plus particuliers, mais la multiplicité des plans complique grandement l'exposé et la synthèse. Notons à ce propos que l'analyse factorielle, qui a pour but de résumer l'information initiale apportée par les 14 causes, ne peut être réduite ici à moins de 5 à 6 dimensions (88 % de la variance sont expliqués par les 6 premiers axes) car l'enchevêtrement des relations entre causes et pays est très complexe. La présentation de deux plans factoriels seulement ne permet en aucun cas de traduire correctement l'information initiale : *seule la classification est valable*, les graphi-

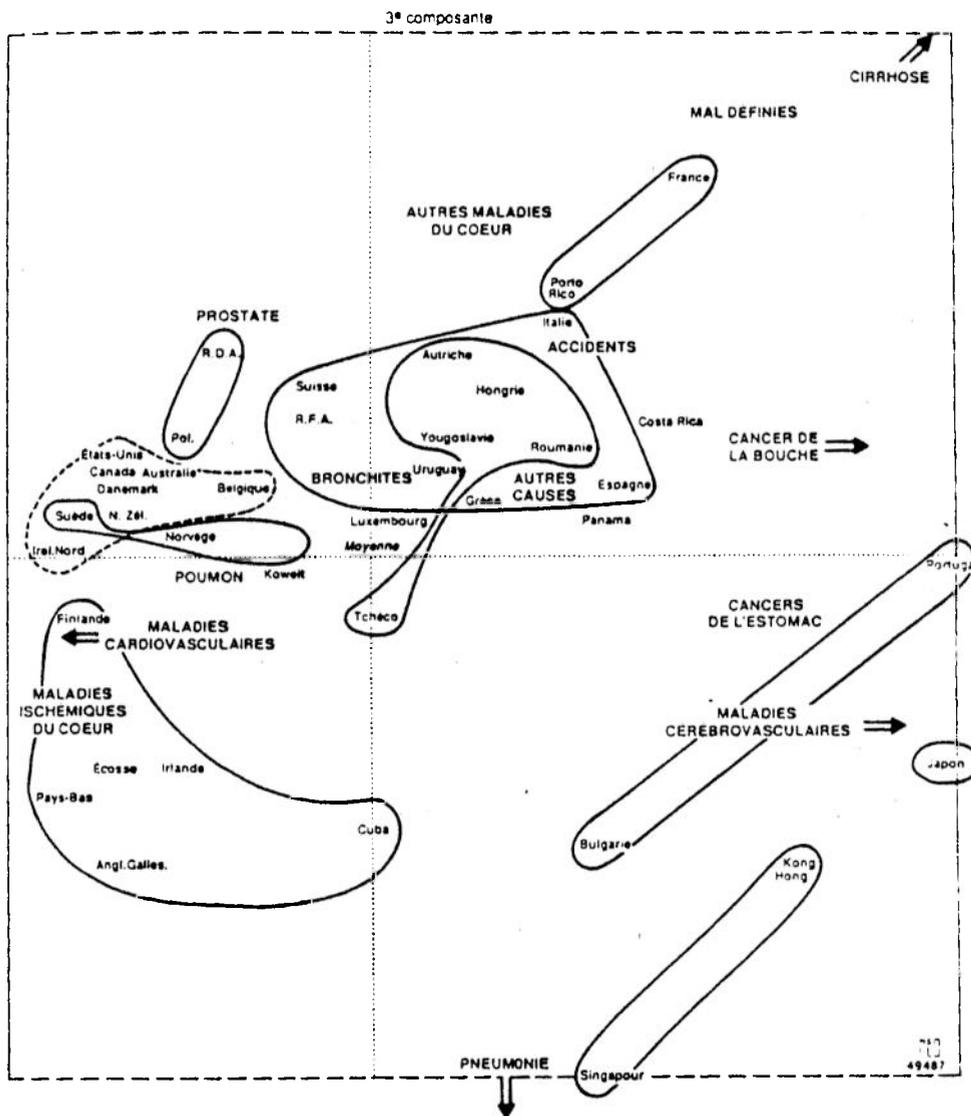


Figure 3. — Classification des pays selon leur profil de décès par cause à 45-54 ans et 65-74 ans. Hommes. Première et troisième composantes

ques présentés ne sont qu'une illustration de cette classification. En effet, si une représentation plane illustre clairement la mortalité par cause, cela signifierait que l'espace des causes pourrait être réduit à deux dimensions ce qui n'aurait pas manqué d'attirer l'attention depuis bien longtemps.

Il est possible d'utiliser les résultats de cette analyse factorielle pour visualiser les positions d'autres pays ou d'un même pays à différentes époques pour voir l'évolution de la répartition par cause indépendamment du niveau. C'est cette technique, classique en analyse des données, que nous avons utilisée pour visualiser les trajectoires des pays développés depuis une quinzaine d'années.

III. — ANALYSE LONGITUDINALE SUR 16 ANS

Certains pays à statistique complète, comme la Suède ou le Japon, disposent de données depuis 1968 jusqu'à 1984, mais d'autres sont plus pauvres. Toutes les données de pays développés ont été projetées sur les différents plans factoriels, mais étant donné que la projection de 640 points sur un même graphique est

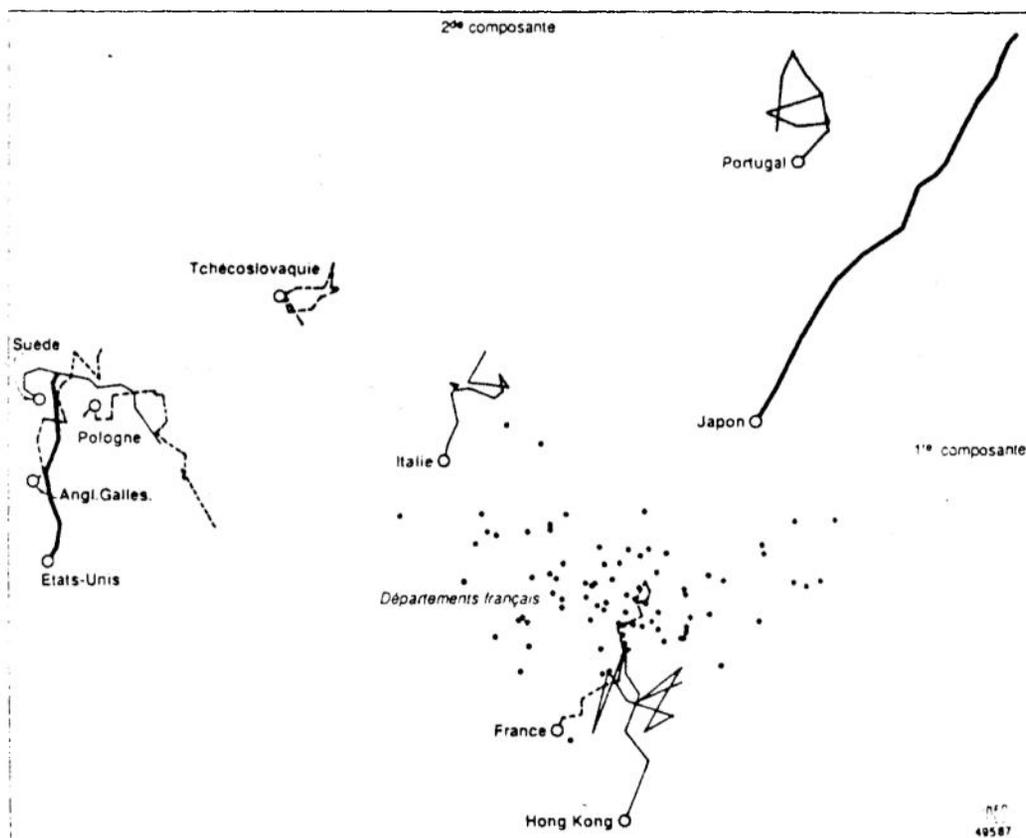


Figure 4. — Trajectoires d'un pays représentant de chaque groupe sur le plan factoriel (1, 2) au cours de la période 1968-1984 et position des 95 départements français

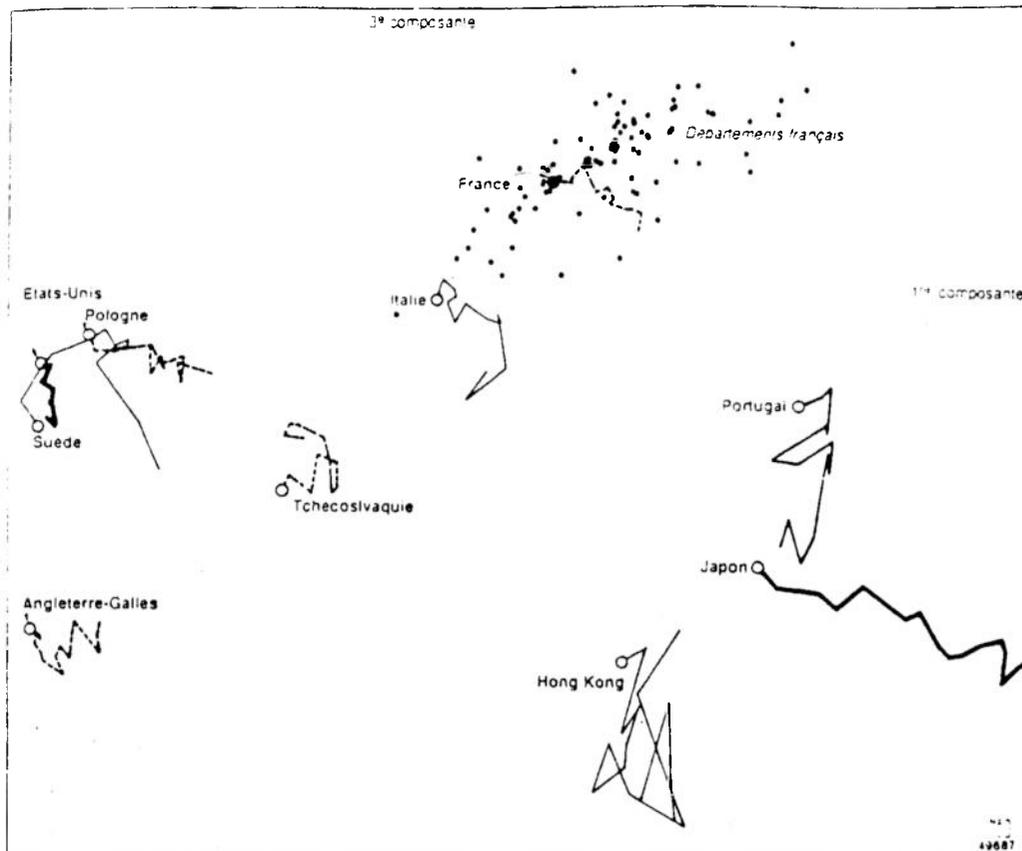


Figure 5. — Trajectoires d'un pays représentant de chaque groupe sur le plan factoriel (1, 3) au cours de la période 1968-1984 et position des 95 départements français

illisible, nous avons choisi de ne faire figurer qu'un représentant de chacun des groupes que nous avons défini.

Les figures 4 et 5 montrent ainsi les trajectoires des pays sur les plans (1,2) et (1,3). A part le Japon, dont la baisse de la mortalité a été particulièrement rapide, nous remarquons qu'en 16 années, les profils de décès par cause évoluent mais restent dans des domaines restreints, ce qui nous permet de dire :

- 1) la typologie proposée résiste à une évolution de moyen terme,
- 2) il n'y a pas d'uniformisation des profils de décès par cause lorsque la mortalité baisse.

Le cas du Japon est intéressant car sa trajectoire montre une baisse très nette des maladies cérébro-vasculaires en plus de la baisse de la mortalité générale (69,05 ans d'espérance de vie pour les hommes en 1968 contre 74,8 ans en 1984 et, respectivement, 74,3 et 80,7 pour les femmes). Les trajectoires inclinées suivant la première bissectrice comme celles de la France ou de l'Italie, montrent aussi une baisse des maladies cérébro-vasculaires.

Pour mieux analyser l'évolution des pays à l'intérieur d'un groupe, nous avons fait des encadrés spécifiques à chaque groupe (voir la figure 6).

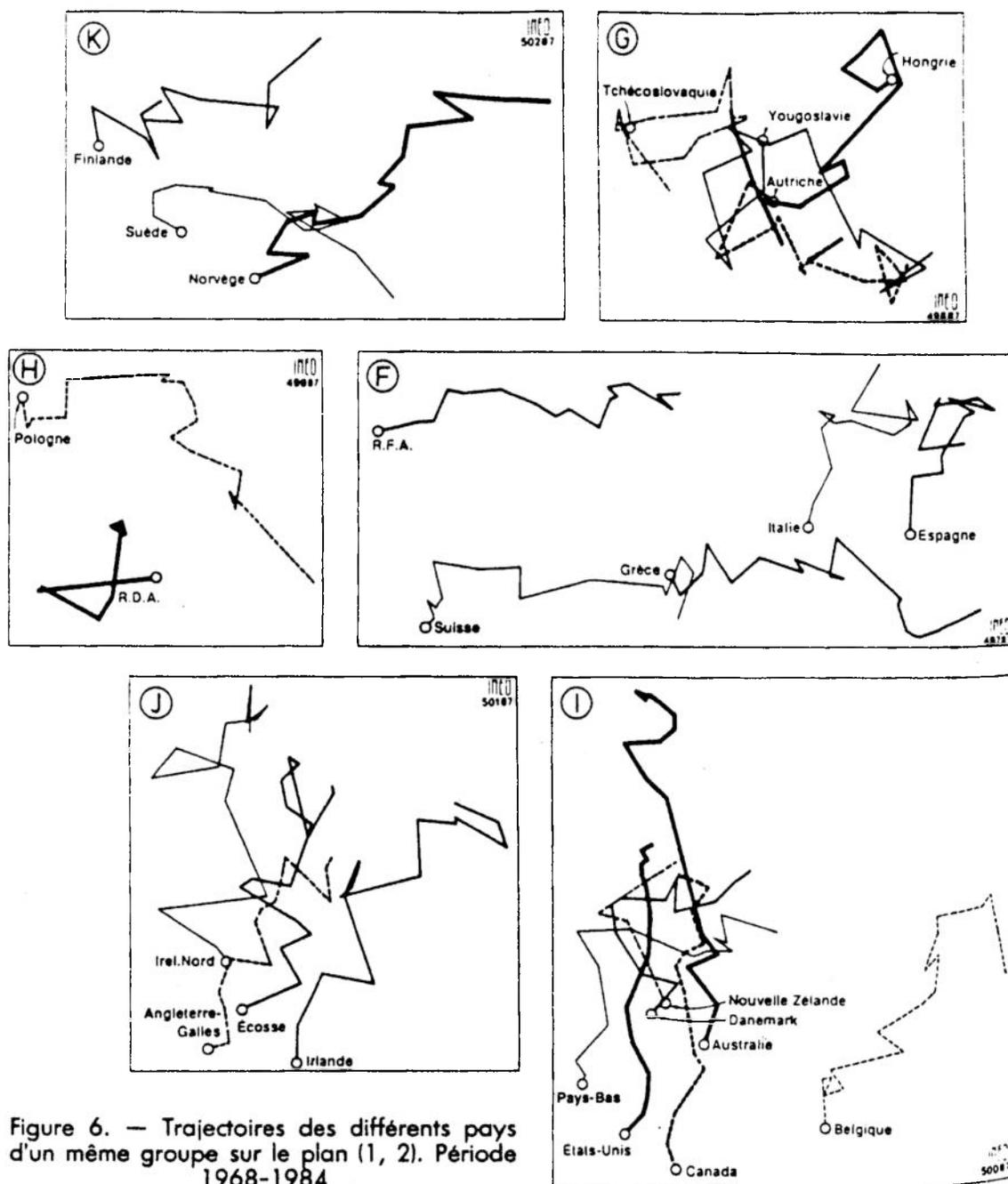


Figure 6. — Trajectoires des différents pays d'un même groupe sur le plan (1, 2). Période 1968-1984

• Dans le groupe « I » des pays anglo-saxons d'outre-mer, associés au Danemark, à la Belgique et aux Pays-Bas, nous remarquons que les trajectoires descendent, montrant par là une baisse de la part des maladies cardio-vasculaires et cérébro-vasculaires au profit d'une augmentation de celle des cancers du poumon.

• L'encadré « J » des îles anglaises montre des tendances tout à fait identiques.

• Celui des pays nordiques « K » est assez différent et il s'agit là d'une augmentation très nette de la proportion de maladies cardio-vasculaires. On remarque le revirement récent de la Suède.

• Dans le groupe « F » des pays méditerranéens, la Grèce diffère de l'Italie et de l'Espagne, par une plus forte proportion de maladies cardio-vasculaires. Les tendances sont identiques pour l'Espagne et l'Italie mais différentes pour la Grèce. La RFA et la Suisse qui étaient regroupées avec ces pays méditerranéens ont tendance à s'éloigner du groupe en raison d'une augmentation régulière de la part des maladies cardio-vasculaires et des cancers du poumon. La RFA se rapproche sensiblement de la RDA.

• Nous en arrivons maintenant au groupe « G » des pays d'Europe centrale : Hongrie, Tchécoslovaquie, Autriche, Yougoslavie et Roumanie. Leur caractéristique principale et très inquiétante est, avant de regarder la répartition par cause, l'augmentation des taux de mortalité masculins à certains âges. Le tableau 1 montre ainsi que, mise à part l'Autriche, les taux de mortalité masculins ont augmenté à tous les âges entre 1968 et 1984. Pour l'Autriche le taux de mortalité à 50 ans a stagné. L'augmentation est moindre à 70 ans, et, pour le sexe féminin la tendance n'est pas systématique.

Notre représentation graphique nous montre que la hausse de la mortalité dans les pays d'Europe centrale n'est pas due à une cause particulière mais concerne toutes les causes. Autrement dit nous assistons à une dégradation générale des conditions de santé dans ces pays. Dans le cas de la Hongrie, on peut remarquer que sa trajectoire remonte selon l'axe 3 (figure 7) ce qui signifie que la mortalité par cirrhose augmente plus que celle attribuable aux autres causes. Ce rôle majeur de l'alcoolisme noté en Hongrie ne se retrouve cependant pas systématiquement dans les autres pays. Il ne caractérise nullement, en particulier, la hausse de la mortalité en Tchécoslovaquie. Des problèmes plus généraux se posent donc, que ne suffit pas à expliquer le contexte politique lui-même, puisque

TABLEAU 1. — TAUX DE MORTALITÉ A 45-54 ET 65-74 ANS POUR LES PAYS D'EUROPE CENTRALE POUR QUI LA MORTALITÉ DES HOMMES A STAGNÉ OU AUGMENTÉ ENTRE 1968 ET 1984

Pays	Année	Hommes		Femmes	
		45-54	65-74	45-54	65-74
Hongrie	1969	72	520	43	339
	1984	133	589	53	335
Tchécoslovaquie	1968	77	542	41	318
	1983	103	606	42	342
Roumanie	1969	70	496	46	360
	1983	93	495	45	334
Yougoslavie	1968	70	470	44	340
	1982	86	500	40	314
Bulgarie	1968	59	411	36	311
	1983	80	509	37	325
Pologne	1969	77	507	42	302
	1984	108	514	41	284
Autriche	1969	77	582	44	321
	1984	76	437	33	231

Source : OMS

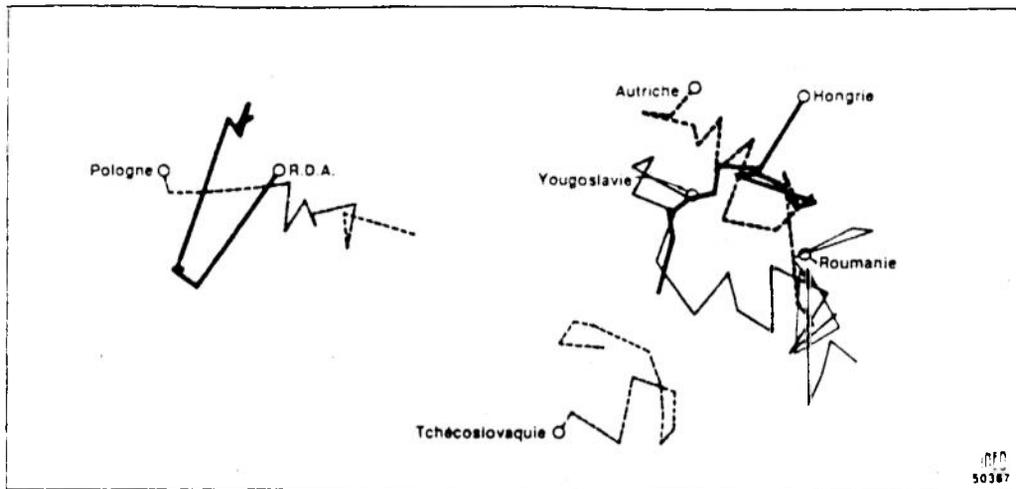


Figure 7. — Hausse de la mortalité dans les pays d'Europe centrale durant la période 1968-1984. Visualisation sur les axes (1, 3) : la trajectoire de la Hongrie s'oriente vers le quart Nord-Est des cirrhoses du foie

la situation des autrichiens n'est pas non plus très favorable. Il semble qu'il y ait actuellement un problème spécifique aux hommes d'Europe centrale.

- Le tableau 1 indique aussi une hausse de la mortalité en Bulgarie et en Pologne, pays qui avec la RDA font partie du groupe « H ». La mortalité baisse en RDA mais la hausse de la mortalité en Pologne est pour l'essentiel imputable aux maladies cardio-vasculaires et aux cancers du poumon.

Ce mode de représentation graphique permet de déceler les grandes tendances de la mortalité par cause, mais ne se substitue pas au travail quantitatif classique qui exige le retour aux données de base. Nous pouvons remarquer aussi que pour un petit pays, ou pour un pays dont les statistiques sont peu fiables, les trajectoires sont très chaotiques : cette représentation permet alors de porter un jugement sur la qualité des données.

Dans le même ordre d'idées, nous nous sommes demandé si la classification présentée ici résistait à la prise en compte des variations géographiques au sein d'un pays. Pour pouvoir en juger nous avons projeté les données relatives aux 95 départements français de la période 1974-1976, sur ces plans factoriels. Le nuage ainsi formé apparaît sur les figures 4 et 5 : un algorithme de classification automatique classe les 95 départements dans le même groupe, mais la dispersion n'est pas négligeable.

IV. — LES RISQUES DE DÉCÈS À 15-24 ANS

Les deux groupes d'âges que nous avons choisis, 45-54 ans et 65-74 ans, concernent la plupart des décès, mais la mortalité à 15-24 ans a fait l'objet d'un intérêt tout particulier ces dernières années du fait de son augmentation dans certains pays. La figure 8 rappelle la position de tous les pays développés vis-à-vis

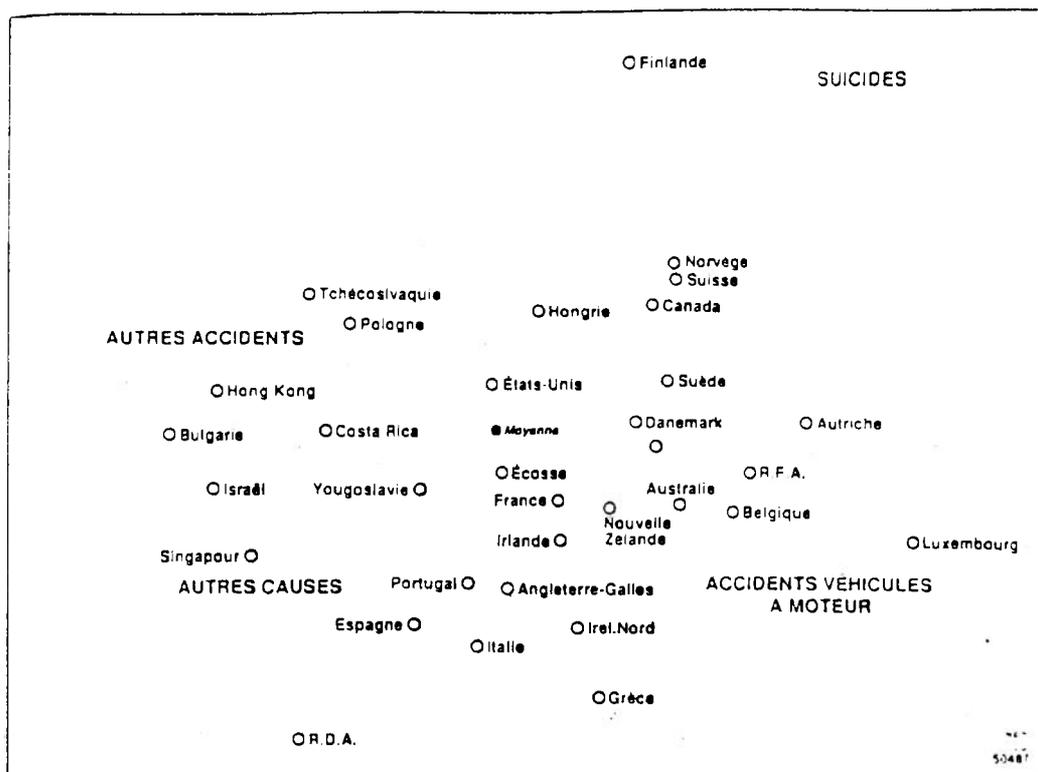


Figure 8. — Position statique des pays en fonction des causes de décès à 15-24 ans. Hommes

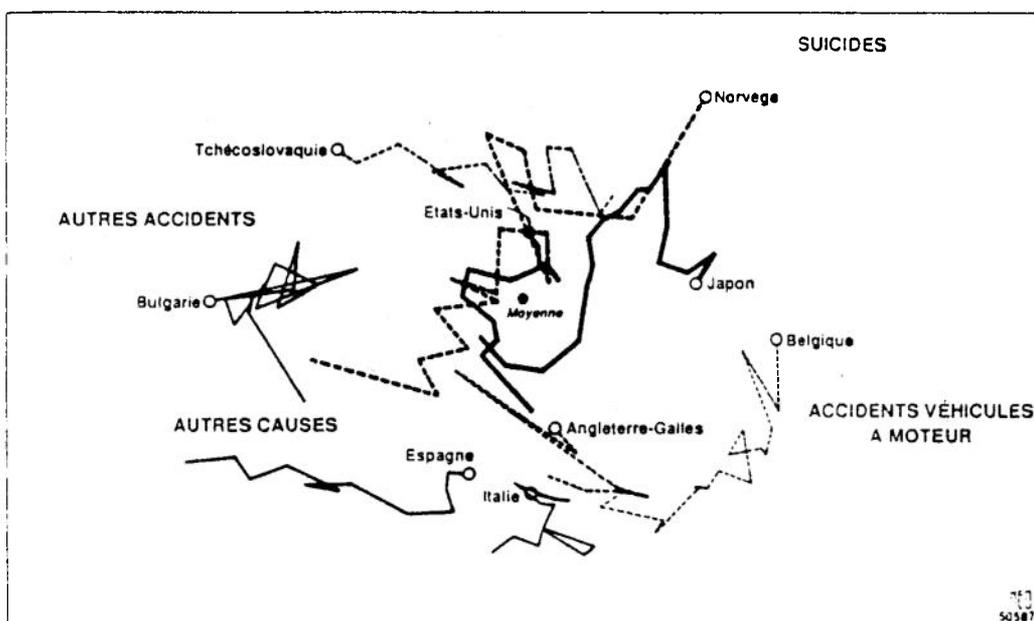


Figure 9. — Trajectoires de quelques pays dans l'espace des causes à 15-24 ans. Hommes

des causes de décès à cet âge. La figure 9 illustre quelques trajectoires nationales remarquables. Ainsi en Norvège la proportion des suicides a pris une place grandissante. La Belgique est toujours particulièrement exposée aux accidents de véhicule à moteur. L'Italie et surtout l'Espagne quittent les profils caractéristiques des pays peu développés (proportion faible de suicides et d'accidents de véhicule à moteur) pour rejoindre les autres pays développés. La Tchécoslovaquie s'éloigne de ce groupe. La Bulgarie stagne.

Notre but était ici de valider une typologie en la soumettant au test d'une analyse longitudinale d'une quinzaine d'années. La relative bonne tenue de cette typologie ne signifie pas que celle-ci soit unique ni optimale; par contre il nous paraît certain que les caractéristiques d'un pays ont une inertie temporelle importante. La construction de tables types de mortalité par cause n'est peut-être pas une utopie.

ANNEXE 1

TABLEAU A. — NUMÉROS DES CAUSES SÉLECTIONNÉES ET CORRESPONDANCE ENTRE LA 8^e ET LA 9^e RÉVISION

Liste des causes	Rubriques de la liste A de la 8 ^e révision	Rubriques de la la 9 ^e révision
Toutes causes	A00	B00
Cancer du poumon	A51	101
Cancer de la bouche, pharynx, œsophage	A45, A46, A50	08, 090, 100
Cancer du sein	A54	113
Cancer de la prostate	A57	124
Cancer de l'estomac	A47	091
Mal. ischémiques	A83	27
Autres formes de mal. du cœur	A84, A86, A87, A88	28, 30
Mal. cérébro-vasculaires	A85	29
Bronchites, asthmes	A93	32
Pneumonie	A91, A92	321
Cirrhose du foie	A102	347
Acc. de véh. moteur	AE138	E471
Suicides	AE147	E54
Autres accidents	AE139-146, 148-150	E470, E472, E479
Mal définies	AE136-137	46
Autres causes	le reste	le reste

ANNEXE 2 RÉSULTATS DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES

Considérons le tableau M_{50} qui fournit le taux de mortalité du pays i ($i \in [1, n]$) pour la cause k ($k \in [1, c]$) entre 45 et 54 ans ainsi que le tableau M_{70} pour les âges 65-74 ans. Juxtaposons-les l'un à côté de l'autre comme le montre la figure suivante :

$$\text{pays } i \quad \begin{bmatrix} & \text{45-54 ans} & \\ & k & \\ \vdots & \vdots & \vdots \\ \dots & m_{i,k}^{50} & \dots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix} \quad \begin{bmatrix} & \text{65-74 ans} & \\ & k & \\ \vdots & \vdots & \vdots \\ \dots & m_{i,k}^{70} & \dots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Appelons $m_{i,j}$ la matrice complète avec $j \in [1, 2c]$. Alors, $m_{i,\cdot} = \sum_j m_{i,j}$ est la somme pour le pays i des taux à 50 ans et 70 ans. De même $m_{\cdot,j} = \sum_i m_{i,j}$ est égal au facteur n près au taux moyen par cause.

Calculons le profil par cause d'un pays i : $\frac{m_{i,j}}{m_{i,\cdot}}$. Définissons une distance (χ^2) entre les profils de deux pays i et i' :

$$d^2(i, i') = \sum_j \frac{1}{m_{\cdot,j}} (m_{i,j} - m_{i',j})^2$$

La pondération $m_{\cdot,j}$ modère l'importance des causes fréquentes et rétablit un certain équilibre entre les taux à 50 ans et ceux à 70 ans qui sont dans un rapport 5 à 6.

Cette distance $d^2(i, i')$ peut aussi s'écrire :

$$d^2(i, i') = \sum_j \left(\frac{m_{i,j}}{m_{i,\cdot} \sqrt{m_{\cdot,j}}} - \frac{m_{i',j}}{m_{i',\cdot} \sqrt{m_{\cdot,j}}} \right)^2$$

et apparaît comme la distance euclidienne entre deux points i et i' de coordonnées \hat{x}_i et $\hat{x}_{i'}$ dans l'espace des causes à $2c$ dimensions :

$$\hat{x}_{i,j} = \frac{m_{i,j}}{m_{i,\cdot} \sqrt{m_{\cdot,j}}}$$

Si, de plus on donne à chaque pays une masse, $m_{i,\cdot}$ (égale ici à la somme des taux de mortalité à 50 et 70 ans toutes causes réunies), l'analyse factorielle des correspondances n'est qu'une analyse en composantes principales classique du tableau \hat{x}_i , pondérée par les masses $m_{i,\cdot}$. On cherche alors la meilleure représentation possible de ce nuage pondéré des n pays dans l'espace à $2c$ dimensions. Le centre de gravité du nuage a pour $j^{\text{ème}}$ coordonnée : $\bar{x}_j = \sqrt{m_{\cdot,j}}$. Soit \vec{u} un vecteur unitaire de \mathbb{R}^{2c} , notons

$$\psi_i = (\vec{x}_i - \vec{\bar{x}}) \cdot \vec{u} = \sum_j \left(\frac{m_{i,j} - m_{i,\cdot} \bar{x}_j}{m_{i,\cdot} \sqrt{m_{\cdot,j}}} \right)$$

la coordonnée d'un point-pays sur l'axe u . Le vecteur \vec{u} sera vecteur principal d'inertie à condition qu'il minimise la quantité : $\sum_i m_{i,\cdot} \psi_i^2$, ce qui correspond après

simplification à la diagonalisation de la matrice $X'X$ ou X est la matrice de terme général :

$$X_{ij} = (m_{ij} - m_{.j}m_{.i}/\sqrt{m_{.j}m_{.i}}).$$

On montre assez aisément que cette matrice symétrique $X'X$ a les mêmes valeurs propres que la matrice X^*X^* de terme plus simple :

$$X^*_{ij} = m_{ij}/\sqrt{m_{.j}m_{.i}}$$

Le vecteur \bar{x} qui correspond au centre de gravité du nuage des pays est alors vecteur propre associé à la valeur propre 1. A part cette valeur propre singulière, notons \bar{u}_α le vecteur propre d'ordre α ; la coordonnée d'un pays i sur l'axe factoriel d'ordre α est alors :

$$\psi_{\alpha i} = \sum_{j=1}^{2c} \left(\frac{m_{ij}}{m_{.j}\sqrt{m_{.i}}} \right) u_{\alpha j}. \quad (1)$$

Représentation duale On peut montrer⁽⁹⁾ que la matrice X^*X^* a les mêmes valeurs propres non nulles que la matrice $X'X$. Si on note v_α le nouveau vecteur propre attaché à la même valeur propre λ_α , on obtient les relations de dualité suivantes :

$$\begin{cases} v_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} X^* u_\alpha \\ u_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} X' v_\alpha. \end{cases} \quad (2)$$

Dès lors, si au lieu de considérer n points-pays dans un espace à $2c$ dimensions, on considère d'une manière symétrique $2c$ points-causes dans un espace à n dimensions, munis des masses $m_{.j}$, on obtient une relation symétrique de la relation 1 qui donne la coordonnée $\phi_{\alpha j}$ de la $j^{\text{ème}}$ cause sur la $\alpha^{\text{ème}}$ vecteur propre v_α , attaché à la même valeur propre λ_α , soit :

$$\phi_{\alpha j} = \sum_{i=1}^n \left(\frac{m_{ij}}{m_{.i}\sqrt{m_{.j}}} \right) v_{\alpha i}. \quad (3)$$

Les relations 1, 2 et 3 conduisent à des relations barycentriques fondamentales de *correspondance* entre les causes et les pays :

$$\begin{cases} \phi_{\alpha j} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n (m_{ij}/m_{.i}) \psi_{\alpha i}, \\ \psi_{\alpha i} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^{2c} (m_{ij}/m_{.j}) \phi_{\alpha j}, \end{cases} \quad (4)$$

Représentation graphique et résultats

Lorsqu'on superpose sur un même graphique le points-causes et les points-pays, on obtient cette correspondance barycentrique.

Comme dans une analyse en composante principale, le pourcentage d'inertie expliqué par les premiers axes est lié aux valeurs propres. Le tableau B donne les valeurs propres ainsi que les pourcentages d'inertie de chaque axe.

⁽⁹⁾ BENZECRI (J.-P.) — *L'analyse des données...* op. cit., note 7.

TABLEAU B. — POURCENTAGE D'INERTIE DES PREMIERS AXES

Axe	Valeur propre	Pourcentage	Pourc. cumulé
1	0,03903	36,26	36,26
2	0,01926	17,90	54,16
3	0,01515	14,08	68,24
4	0,00967	8,99	77,23
5	0,00771	7,16	84,39
6	0,00423	3,93	88,32
7	0,00393	3,66	91,98
8	0,00318	2,95	94,93
9	0,00196	1,82	96,75

Nous nous sommes limité ici à la représentation des 3 premiers axes (68 % de la variance). Partant des groupes de causes du tableau A, nous avons exclu les causes mal définies⁽¹⁰⁾, et regroupé d'une part tous les accidents et d'autre part toutes les maladies du cœur. Nous avons donc en tout 10 causes pour chaque groupe d'âge, soit 20 variables.

Le tableau C donne pour chacune de ces variables, la masse m_i , et ses coordonnées sur les 5 premiers axes (ϕ_{ij}).

Une fois que le nuage des causes est placé, les pays sont positionnés selon la relation barycentrique 4. Dans la communication présentée au congrès de Florence⁽¹¹⁾, nous avons

TABLEAU C. — COORDONNÉES DES POINTS-CAUSES SUR LES PREMIERS FACTEURS

Cause	Masse m_i	Coordonnées sur les facteurs					
		1	2	3	4	5	6
A 50 ans							
Poumon	0,012	-0,05	-0,02	0,05	-0,21	0,01	0,07
Bouche	0,005	0,45	-0,39	0,13	-0,27	-0,18	0,17
Estomac	0,004	0,34	0,20	-0,07	0,17	0,21	-0,02
Cardio-vasculaire	0,043	-0,21	0,05	-0,04	0,00	-0,09	0,02
Cérébro-vasc.	0,009	0,27	0,18	-0,16	0,01	-0,03	0,05
Bronchites	0,002	0,04	0,13	0,00	-0,03	0,06	0,25
Pneumonie	0,002	0,31	-0,15	-0,35	0,15	-0,20	0,27
Cirrhose	0,008	0,41	0,06	0,35	-0,06	-0,17	0,04
Tous accidents	0,021	0,09	0,06	0,15	0,16	-0,05	0,08
Autres causes	0,038	0,13	-0,09	0,06	0,04	-0,05	0,05
A 70 ans							
Poumon	0,072	-0,17	-0,12	-0,08	-0,22	0,12	-0,02
Bouche	0,017	0,27	-0,34	0,04	-0,08	-0,10	0,09
Estomac	0,028	0,29	0,09	0,00	0,23	0,29	-0,05
Cardio-vasculaire	0,318	-0,20	0,05	0,02	0,04	-0,04	-0,01
Cérébro-vasc.	0,102	0,27	0,26	-0,11	-0,10	-0,04	-0,05
Bronchites	0,034	-0,09	0,15	0,06	-0,07	0,21	0,23
Pneumonie	0,024	0,15	-0,20	-0,58	0,16	-0,08	0,13
Cirrhose	0,016	0,46	0,05	0,37	-0,12	-0,15	0,09
Tous accidents	0,031	0,16	0,00	0,13	0,12	0,09	0,02
Autres causes	0,215	0,08	-0,15	0,03	0,02	0,01	-0,05

(10) Ceci revient à répartir les mal définis proportionnellement aux autres causes.

(11) BROUARD (Nicolas) et LOPEZ (Alan). — Causes of death patterns... op. cit., note 3

sélectionné les données les plus récentes des pays ayant plus de 70 ans d'espérance de vie; ici nous avons conservé cette première sélection (et donc diagonalisé la même matrice) et utilisé ces résultats pour projeter soit d'autres pays, soit un même pays à diverses dates.

Prenons l'exemple d'un département français, dont on connaît les taux de mortalité masculins selon les 10 causes retenues (tableau A) et pour les deux groupes d'âges 45-54 ans et 65-74 ans. Notons t_j , $\{j \in [1, 2 \times 10]\}$, ces taux, et $t = \sum_j t_j$ la somme des 20 taux toutes causes. Alors, la coordonnée, ψ_α , sur l'axe α de ce département est donnée par l'équation :

$$\psi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^{20} (t_j/t) \phi_{\alpha j}. \quad (5)$$

Numériquement, et pour le premier axe, nous obtenons (voir les tableaux B et C) :

$$\psi_\alpha = \frac{1}{\sqrt{0,03903}} (-0,05 \text{ (poumon 50 ans)} + 0,45 \text{ (bouche 50)} + 0,34 \text{ estomac} \quad (6)$$

$$+ 0,27 \text{ cérébro} + \dots + 0,12 \text{ (poumon 70 ans)} + \quad (7)$$

$$\dots + 0,08 \text{ (toutes les autres causes 70 ans)).} \quad (8)$$

En utilisant ce procédé, il est possible aussi de visualiser d'autres causes ou sous-causes. C'est ce que nous avons fait sur la figure 2 en éclatant l'ensemble des maladies cardiovasculaires en « Maladies ischémiques du cœur » et « Autres maladies du cœur ». Pour plus de détail sur la programmation on peut se référer à l'ouvrage⁽¹²⁾

⁽¹²⁾ LEBART (L.), MORINEAU (A.) et FÉNÉLON (J.P.). — *Traitement des données statistiques Méthodes et programmes*. — Paris, Dunod, 1980, 512 p.