

# JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

ANDRÉE MIZRAHI

ARIÉ MIZRAHI

## **Qualité des données et méthodes d'observation**

*Journal de la société statistique de Paris*, tome 124, n° 1 (1983), p. 19-37

[http://www.numdam.org/item?id=JSFS\\_1983\\_\\_124\\_1\\_19\\_0](http://www.numdam.org/item?id=JSFS_1983__124_1_19_0)

© Société de statistique de Paris, 1983, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

# QUALITÉ DES DONNÉES ET MÉTHODES D'OBSERVATION (1)

Andrée et Arié MIZRAHI

*Division d'économie médicale du C.R.E.D.O.C.*

*En vue de mettre au point les conditions optimales d'une enquête auprès des médecins, sur la morbidité et la thérapeutique médicale, une enquête expérimentale a été effectuée en juin-juillet 1980 auprès d'un échantillon de 582 médecins, qui devaient remplir une fiche pour chaque patient vu pendant une période de deux ou sept jours.*

*Sept variables contrôlées sont étudiées :*

- le mode et la nature du contact.*
- la durée de la période d'observation auprès de chaque médecin enquêté,*
- la rémunération du médecin enquêté, la nature de l'enquêteur, la longueur du questionnaire patient, la société d'enquête.*

*Le contact épistolaire entraîne une participation si faible qu'elle s'apparente au volontariat; l'appel téléphonique induit une meilleure participation des médecins que la visite de l'enquêteur, sans doute par les facilités d'encadrement qu'il offre; la modalité la plus favorable, de ce point de vue, consiste à téléphoner sans avoir demandé au préalable au médecin de se déterminer. L'allongement de la durée d'observation diminue la participation, alors qu'une rémunération l'augmente sensiblement. Le questionnaire le plus bref, dont les questions sont les plus simples, est le mieux accueilli, le plus complet le moins bien. Enfin, l'intervention de deux sociétés met en évidence le rôle important de l'encadrement des enquêteurs. De même, la quantité d'informations recueillies, et sa précision, dépend des modalités de l'enquête; la durée par exemple est une cause de dégradation grave, puisque, en moyenne, chaque jour additionnel, les médecins fournissent environ 13 à 14 % de moins de fiches séances que la veille; le contenu de ces séances (nombre d'ordonnances, nombre de lignes d'ordonnance) est peu modifié, mais le nombre de diagnostics décroît également.*

*Les modalités de chaque enquête dépendent de la nature des recherches qui en seront issues et de l'importance qu'on accorde aux différents biais.*

*In order to define the optimal conditions of a survey among physicians on morbidity and medical therapeutics, an experimental survey has been carried out in June-July 1980, on a sample of 582 physicians, who were to fill out a form for each patient examined over a period of two to seven days.*

*Seven controlled variables have been studied: mode and nature of the contact, duration of the observation period for each physician surveyed, remuneration of the surveyed doctor, nature of the surveyor, length of the patient-questionnaire, the survey society.*

## PRÉSENTATION

Trois types de problèmes se posent dès qu'on cherche à analyser la qualité des données obtenues par enquête aléatoire dans le domaine des sciences sociales :

- La représentativité de l'échantillon, et les problèmes entraînés par les refus de répondre à l'enquête, l'abandon lorsque la participation entraîne un travail trop important, les sujets que l'enquêteur n'a pu joindre (absents de longue durée).

---

1. Communication faite le 24 mars 1982 devant les Sociétés de statistique de Paris et de France.

— L'exactitude des réponses des enquêtés, et les biais dus aux oublis, négligences et omissions volontaires, ainsi qu'à l'incompréhension de certaines questions.

— La perte de précision sur données individuelles liée à la période d'enquête, et l'accroissement de la variance d'échantillonnage lorsque cette période est raccourcie.

Ces trois problèmes sont intimement liés : un allongement de la période d'enquête réduira la variance d'échantillonnage mais entraînera, par usure, un plus grand taux d'abandon et une diminution de l'exactitude des réponses; une pression trop forte exercée sur les récalcitrants pour participer en plus grand nombre à l'enquête peut entraîner une augmentation des négligences et omissions volontaires...

Ces trois points ont pu être abordés à l'occasion d'une enquête expérimentale effectuée en juin-juillet 1980 par le C.R.E.D.O.C. auprès de 582 médecins (<sup>1</sup>); il était demandé à chaque médecin de :

- remplir un questionnaire le caractérisant — temps nécessaire : 15 à 20 minutes;
- fournir, pour chaque patient vu pendant une période à déterminer (variable contrôlée), différentes informations démographiques, sociales, médicales, et sur les thérapeutiques prescrites — temps nécessaire : une à trois minutes par patient.

Le but de l'opération était d'analyser l'influence des différentes variables mesurables sur la qualité des données recueillies.

## 1. REPRÉSENTATIVITÉ DE L'ÉCHANTILLON ET TAUX DE PARTICIPATION

Il ne s'agit pas ici de considérations sur la théorie des sondages : disposant d'un échantillon représentatif tiré à partir d'une bonne base de sondage, on constate qu'une partie des individus, qui aurait dû y participer, échappe à l'enquête. Il s'agit d'absents que l'enquêteur ne peut joindre malgré plusieurs tentatives, ou de personnes refusant de répondre aux questionnaires (ou à une partie importante d'entre eux). L'erreur introduite tient au lien entre la variable étudiée et la participation à l'enquête : les personnes pour lesquelles répondre entraîne le plus gros travail sont celles qui refusent ou abandonnent l'enquête le plus fréquemment.

Le taux de participation dépend de l'âge et du sexe du médecin, de sa nature (généraliste ou spécialiste), de son mode d'exercice (salarié ou libéral, individuel ou de groupe) et de l'importance de son activité, de la taille de sa commune d'exercice.

En l'absence d'information préalable sur les taux de refus, et en partant donc d'un tirage aléatoire, l'échantillon de participants ne peut être que biaisé.

Pour améliorer les résultats, deux voies doivent être poursuivies simultanément :

- accroître le taux de participation à l'enquête en simplifiant et réduisant autant que possible le travail demandé, par une campagne d'information sur les buts poursuivis, et en exprimant aux enquêtés la reconnaissance des responsables de l'enquête (lettres de remerciement, attribution de rémunération, retour d'information, ...);
- chercher à obtenir des renseignements, même rudimentaires, sur les non-participants, et utiliser ces données pour redresser l'information brute.

### 1.1. *Les variables d'action*

Sept variables d'action ont été testées au cours de l'enquête expérimentale auprès des médecins :

- le mode de premier contact (épistolaire ou contact personnel);

1. Cf. Méthode d'enquête : Morbidité et thérapeutique médicale, Ph. LE FUR, Andrée MIZRAHI, Arié MIZRAHI, C.R.E.D.O.C. 1981, 231 pages.

- la nature du contact (appel téléphonique ou visite d'enquêteur);
- la nature de l'enquêteur (médecin ou non-médecin);
- la longueur du questionnaire patient (3 modalités : A, B, C);
- la durée de l'observation auprès de chaque médecin (2 jours, 7 jours);
- la rémunération des médecins enquêtés;
- la qualité de la société d'enquête (2 sociétés testées, S 1 et S 2).

Toutes les modalités élémentaires n'ont pas été testées (ainsi, par exemple, seule une moitié des médecins faisant l'enquête pendant 7 jours se voyait proposer une rémunération), et, au total, la comparaison a porté sur 144 modalités.

Sous réserve d'hypothèses simples et robustes permettant une transformation fonctionnelle inspirée des probits, on montre qu'il est légitime d'admettre l'indépendance de l'effet des différentes variables d'action sur les taux d'acceptation, et d'étudier ainsi successivement chacune de ces variables.

1.1.1. *Mode de contact*

329 médecins ont reçu un premier courrier leur demandant de bien vouloir participer à l'enquête et de le manifester en renvoyant une carte réponse dite carte T (cf. tableau 1); 67 l'ont fait, 4 pour refuser, 63 pour accepter à ce premier stade. Parmi eux, 7 ont abandonné, et 56 ont mené l'enquête jusqu'à son terme (renvoi des documents), soit un taux de participation de  $56/329 = 17,0 \%$  pour la première modalité. Les 262 médecins n'ayant pas répondu ont été contactés par des enquêteurs, selon diverses modalités; au cours de cette relance, on observe 156 refus et 106 acceptations, dont 37 non suivies d'effets (abandons) et 69 participations : la deuxième modalité consistant à recueillir un accord épistolaire et à relancer les non-réponses, conduit donc à un taux global de participation de  $125/329 = 38,0 \%$ .

TABLEAU 1

*Mode de contact et participation*

Modalités	Dans le champ	Refus	Abandon	Participation	Taux de participation	Non réponse
1) Carte T . . . . .	329	4	7	56	17,02 %	262
Résidu Carte T . . . . .	262	156	37	69	—	—
2) Carte T puis entretien	329	160	44	125	37,99 %	—
3) Entretien . . . . .	253	92	54	107	42,29 %	—
Ensemble . . . . .	582	252	98	232	39,86 %	—



La troisième modalité (lettre seulement introductive, l'engagement du médecin étant obtenu par un entretien) conduit à un taux d'acceptation de 42,3 %.

Au total, l'enquête par correspondance est à proscrire, car elle aboutit à un échantillon de volontaires plutôt qu'à un sondage aléatoire; la modalité 2, qui présente l'intérêt de ne contacter par entretien que 79,6 % de l'échantillon de départ (262/329), fournit cependant un taux de participation légèrement inférieur.

1.1.2. *Nature du contact*

En dehors de ceux faisant l'enquête par correspondance, les médecins contactés par l'enquêteur l'étaient, soit par appel téléphonique seulement, soit par visite d'enquêteur (cf. tableau 2). La différence de participation entre les deux modalités est faible (3,32 %). Les taux d'abandons (20,68 % et 13,64 %) ne sont, de même, pas très éloignés.

Deux observations sont à souligner :

- le tableau 2 exclut les médecins ayant répondu à la carte T, ce qui explique les nombres et le taux d'ensemble inférieurs à ceux du tableau 1. La dernière colonne du tableau présente les taux corrigés (il s'agit des résultats obtenus en reventilant les médecins ayant répondu à la carte T).

TABLEAU 2

*Nature du contact et participation*

Modalités	Dans le champ	Refus	Abandon	Participation		
				Effectifs	Taux	Taux <sup>(1)</sup> corrigés
Téléphone . . . . .	295	129	61	105	35,59 %	41,27 %
Visite . . . . .	220	119	30	71	32,27 %	38,00 %
Ensemble . . . . .	515	248	91	176	34,17 %	39,86 %

1. Pour tenir compte du fait qu'une partie des contacts sont des résidus (médecins n'ayant pas répondu aux cartes T).

- les appels téléphoniques sont plus aisés à faire, et peut-être d'un meilleur rapport pour l'enquêteur, que la visite au médecin; en particulier l'encadrement de l'une des deux sociétés ayant un moment faibli, la perte s'est davantage portée sur cette modalité (visite d'enquêteur).

Au total, la visite d'enquêteur est de beaucoup plus onéreuse que l'appel téléphonique qui, de plus, peut être centralisé, ce qui permet un meilleur contrôle du travail : cette modalité apparaît la plus efficace à ce niveau de l'analyse.

1.1.3. *Mode et nature du contact*

Croisant les deux modalités déjà étudiées, il vient cinq modalités combinées (cf. tableau 3). Seul, le contact épistolaire est invariant; les deux modalités avec une carte T préalable donnent des résultats inférieurs, et, de même, celles impliquant un déplacement d'enquêteur sont moins favorables que le contact téléphonique; les résultats sont cependant assez proches, et une seule modalité (carte T puis visite d'enquêteur) s'en écarte de manière importante : elle est liée à la difficulté de bien contrôler en permanence le travail des enquêteurs sur le terrain.

Ces résultats portent sur des effectifs réduits, et on leur associera des estimations établies sous une hypothèse d'indépendance, à partir des résultats obtenus pour chacune des deux modalités — mode de contact — nature du contact — prise isolément (cf. tableau 3, dernière colonne), la modalité croisée la plus efficace consiste à téléphoner sans avoir demandé au préalable au médecin de se déterminer : on obtient, dans ce cas, un taux d'acceptation de 43,71 % (taux estimé).

TABLEAU 3

*Mode et nature du contact et participation*

	Dans le champ effectif	Taux				Taux de participation estimé
		Refus	Abandon	Participation	Non réponse	
Carte T . . . . .	329	1,22	2,13	17,02	79,64	—
Carte T plus appel téléphonique	153	44,94	14,62	40,44	—	39,36
Carte T plus visite enquêteur	109	53,83	11,63	34,55	—	36,24
Appel téléphonique . . . . .	142	31,69	26,06	42,25	—	43,71
Visite enquêteur . . . . .	111	42,34	15,32	42,34	—	40,57
Ensemble . . . . .	582	43,30	16,84	39,86	—	—

1.1.4. *Durée de l'enquête auprès de chaque médecin*

Une partie des médecins (190) était priée de fournir des informations sur une période de deux jours, l'autre (392) pendant une semaine; les taux de participation ne sont pas très différents selon ces deux modalités (cf. tableau 4); mais cela est dû au fait qu'une partie de l'échantillon « semaine » était rémunéré. Il est vraisemblable qu'en rémunérant l'échantillon deux jours, on obtiendrait de bien meilleurs résultats (cf. § 116 — Durée d'enquête et rémunération).

TABLEAU 4

*Durée de l'enquête et participation*

	Dans le champ	Refus	Abandon	Participation	Taux de participation
2 jours . . . . .	190	79	29	82	43,16
1 semaine . . . . .	392	173	69	150	38,27
Ensemble . . . . .	582	252	98	232	39,86

1.1.5. *Rémunération*

Pour 165 médecins sur les 582 dans le champ, une rémunération était proposée en échange du service rendu. Les taux de participation sont très sensiblement différents (cf. tableau 5); pourtant, les médecins rémunérés appartenaient tous à l'échantillon 7 jours; ici aussi, il est vraisemblable que si l'échantillon rémunéré avait été ventilé en 2 jours — 1 semaine, on aurait observé un taux de participation supérieur à 55,15 % (cf. § 116 — Durée d'enquête et rémunération).

TABLEAU 5

*Rémunération et participation*

	Dans de champ	Refus	Abandon	Participation	Taux de participation
Rémunération . . . . .	165	57	17	91	55,15 %
Pas de rémunération . . . . .	417	195	81	141	33,81 %
Ensemble . . . . .	582	252	98	232	39,86 %

1.1.6. *Durée d'enquête et rémunération*

L'échantillon global des médecins était ventilé en trois groupes :

- 2 jours et non rémunéré (2 J R),
- 7 jours et rémunéré (7 J R),
- 7 jours et non rémunéré (7 J R).

Il n'y avait pas de modalité 2 jours et rémunéré (2 J R).

La participation est bien meilleure (55,2 %) lorsqu'on rémunère les médecins, même pour une période relativement longue (une semaine) que lorsqu'on réduit cette période à 2 jours sans rémunération.

L'hypothèse d'indépendance de la variable transformée (probit) permet d'estimer, connaissant les trois modalités 2 J R, 7 J R, 7 J R, le taux de participation qu'on aurait observé pour un échantillon de médecins rémunérés observés pendant 2 jours (2 J R) : on obtiendrait un taux de 72,60 % d'acceptations (cf. tableau 6).

TABLEAU 6

*Durée d'enquête, rémunération et participation*

	Dans le champ	Refus	Abandon	Participation	Taux de participation	
					observé	estimé
2 J R	—	—	—	—	—	72,60
2 J R . . . . .	190	79	29	82	43,16	—
7 J R . . . . .	165	57	17	91	55,15	—
7 J R . . . . .	227	116	52	59	25,99	—
Ensemble	582	252	98	232	39,86	—

1.1.7. *Autres modalités contrôlées. (Type d'enquêteur, modèle de questionnaire, société d'enquête)*

L'enquêteur, lorsqu'il est lui-même médecin a, en moyenne, un meilleur résultat sur les taux d'acceptation, que l'enquêteur non médecin; sans être très important (cf. 1<sup>re</sup> partie du tableau 7) l'effet n'est toutefois pas négligeable.

Le modèle de questionnaire (de difficulté croissante, A, B, C) a également un léger effet sur les taux d'acceptation. Les meilleurs résultats sont obtenus par le questionnaire A (cf. 2<sup>e</sup> partie du tableau 7), et le questionnaire B donne des résultats légèrement inférieurs (3,35 % de moins, en taux corrigés) <sup>(1)</sup>.

1. Si la ligne C du tableau 7, 2<sup>e</sup> partie, n'est pas comparable aux deux premières, car ne comportant que 2 J et 7 J R, les lignes A et B le sont tout à fait entre elles; or, les taux de refus sont proches dans ces deux cas (respectivement 44 % et 44,7 %), mais le taux d'abandon, parmi les médecins ayant accepté (36 + 80 = 116 et 40 + 70 = 110), est supérieur pour B (36,4 %) à celui de A (31 %); cet effet traduit peut être une certaine lassitude devant la difficulté plus grande de B.

TABLEAU 7

*Type d'enquêteur, modèle de questionnaire, société d'enquête et participation*

Modalités	Dans le champ	Refus	Abandon	Participation	Taux de participation	
					observé	corrigé (1)
<i>Enquêteur</i>						
Médecin . . . . .	263	111	56	96	36,50	42,03
Non médecin . . . . .	252	137	35	80	31,75	38,03
Correspondancé . . . . .	329	4	7	56	17,02	—
<i>Questionnaire</i>						
A . . . . .	207	91	36	80	38,65	43,09
B . . . . .	199	89	40	70	35,16	39,74
C . . . . .	176	72	22	82	46,59	37,03
<i>Société</i>						
S 1 . . . . .	278	109	42	127	45,68	
S 2 . . . . .	304	143	56	105	34,54	
<b>Ensemble</b> . . . . .	<b>582</b>	<b>252</b>	<b>98</b>	<b>232</b>	<b>39,86</b>	

1. Pour tenir compte du fait qu'une partie des contacts sont des résidus (médecins n'ayant pas renvoyé les cartes T) et que le questionnaire C n'était pas adressé à l'échantillon 7 J R

Enfin, le questionnaire C (2,71 % de moins que B, et 6,05 % de moins que A en taux corrigés) cumulant les difficultés, donne le taux de participation (corrigé le plus bas).

La dernière partie du tableau 7 montre l'importance d'un contrôle minutieux et permanent du réseau d'enquêteurs; en effet, la personne chargée du réseau d'enquêteurs dans S 2 a quitté l'entreprise sans être immédiatement remplacée, le responsable direct de cette personne étant lui-même absent pour quelques jours; jusqu'au remplacement de la personne démissionnaire, les résultats de S 2 se sont gravement dégradés, sans que les mesures prises par la suite puissent rattraper ce handicap. Le tableau 8 présente les résultats des deux sociétés selon le mode de contact; or, pour des raisons matérielles (retour des cartes T), les modalités avec et sans carte T se sont déroulées à des moments différents. La différence de résultats entre S 1 et S 2 est nettement plus grave pour la modalité sans carte T que pour celle avec carte T, car c'est au moment où la modalité sans carte T se déroulait que l'absence de la personne chargée du réseau d'enquêteurs s'est fait le plus sentir.

Du fait du déroulement dans le temps des modes de contact, et des difficultés de S 2, les variables mode de contact et société d'enquête ne sont pas indépendantes, et les effectifs estimés dans le cas d'indépendance (dernière colonne du tableau 8) sont éloignés des effectifs observés. Cette liaison étant conjoncturelle, nous n'en tiendrons pas compte, et nous régulariserons les résultats de S 2, sachant que nous pénalisons pour cette société la modalité avec carte T, mais qu'en échange, nous avantageons la modalité sans carte T. Une telle régularisation a un effet plus pernicieux sur la comparaison des modes de contact : l'écart entre les deux modes testés est écrasé, puisqu'il passe de 40,63 % à 52,54 % pour S 1. Dans l'interprétation des résultats, il faudra tenir compte du fait que l'hypothèse d'indépendance réduit l'écart entre les deux modes de contact, et que la modalité sans carte T est plus favorable que les résultats globaux ne le laissent paraître.



TABLEAU 8

*Société d'enquête, mode de contact et participation*

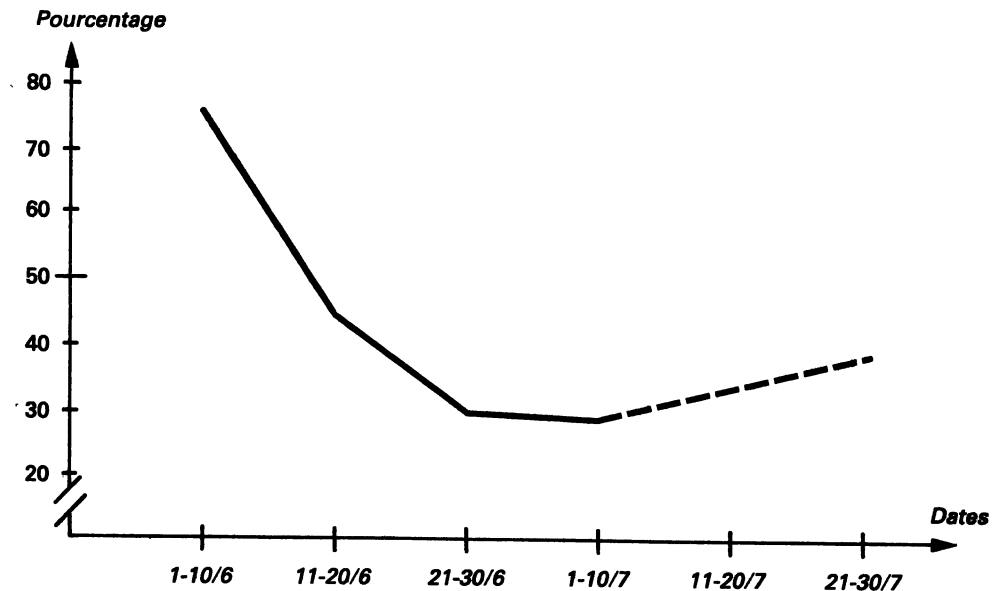
Modalités		Dans le champ	Participation		Effectif estimé
Société	Contact		Effectif	Taux	
S 1	avec carte T	160	65	40,63	70
	sans carte T	118	62	52,54	57
S 2	avec carte T	169	60	35,50	55
	sans carte T	135	45	33,33	50

1.2. *Les autres variables*1.2.1. *Variables décrivant l'exécution de l'enquête*

Ce sont des variables liées à l'enquête expérimentale de juin-juillet 1980, et vraisemblablement non reproductibles : il s'agit de la date du premier contact, de l'existence d'un rappel la veille de l'enquête, du nombre d'entretiens et de leur durée totale.

La date du premier contact est liée aux modalités de l'enquête : Les premiers médecins avaient seulement un contact épistolaire (renvoi de la carte T), les médecins suivants appartenaient à l'échantillon contact direct sans carte T, les derniers étaient ceux n'ayant pas renvoyé la carte T (cf. graphique 1).

GRAPHIQUE 1

*Date du premier contact et participation*

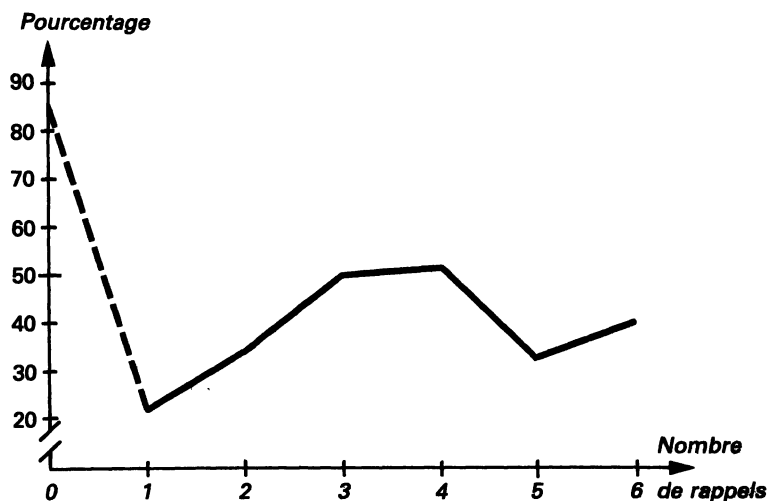
Le rappel la veille de l'enquête devait avoir lieu si les relevés étaient commencés plus de 48 heures après l'acceptation du médecin : sans que l'écart soit significatif, le sens de la différence indique qu'il vaut mieux faire effectuer l'enquête aux médecins aussitôt qu'ils ont exprimé leur accord

(75,5 % de confirmation), car un délai, même de quelques jours, peut entraîner une modification défavorable de l'attitude de quelques-uns d'entre eux (65,5 % de confirmation); on ne peut cependant pas écarter l'idée que certains médecins repoussent le début de l'enquête en raison de réticences inavouées les conduisant, en fin de compte, à une non participation.

Le nombre d'entretiens et leur durée totale sont liés et leurs effets sur la participation (*cf.* graphiques 2 et 3) sont similaires : le zéro correspond à des médecins ayant renvoyé leur carte T (fort taux de participation) et, en dehors de ce point, le taux de participation croît avec l'effort consenti; quoique cette croissance se réduise lorsque le nombre et la durée augmentent, la ténacité est payante et doit être pratiquée systématiquement.

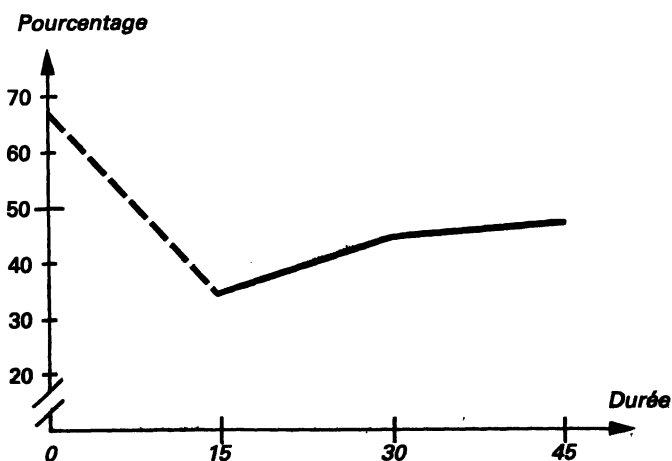
GRAPHIQUE 2

*Nombre d'entretiens et participation*



GRAPHIQUE 3

*Durée totale des entretiens et participation*

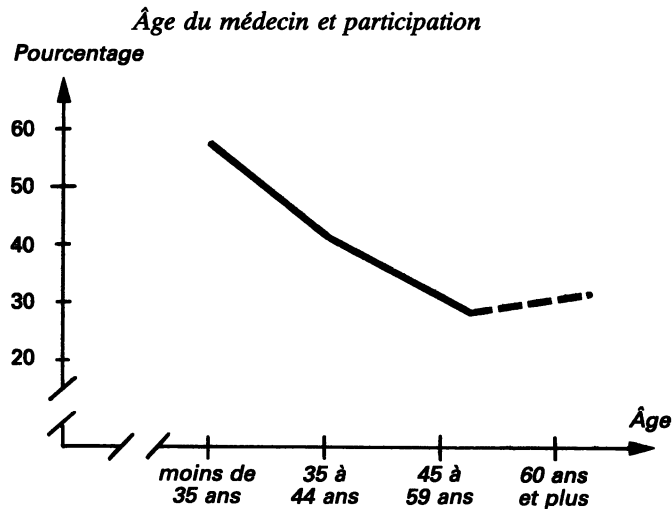


### 1.2.2. Variables décrivant la population enquêtée

Les généralistes (45 %) sont plus coopératifs que les spécialistes (32,3 %), l'écart est important et significatif; les spécialistes ont souvent le sentiment que l'enquête ne les concerne pas vraiment, et la publication des premiers résultats (pathologie rencontrée, actes pratiqués, prescriptions) devrait les convaincre du contraire, et les inciter à participer à l'enquête en plus grand nombre.

les jeunes médecins participent plus souvent que leurs aînés, moins familiarisés aux méthodes de sondage (cf. graphique 4); tout en faisant un effort de vulgarisation des résultats, on peut espérer que les taux de participation iront en s'améliorant avec l'arrivée en plus grand nombre de jeunes médecins.

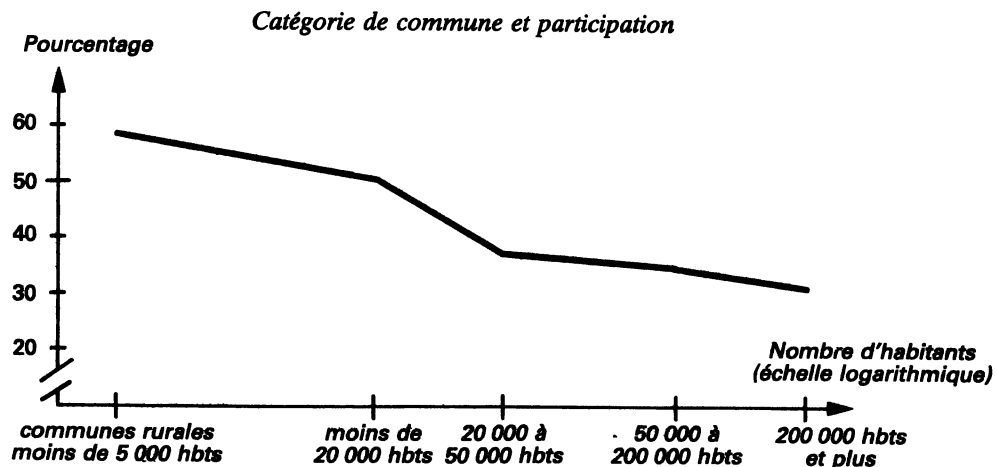
GRAPHIQUE 4



Enfin, l'écart des taux de participation entre les deux sexes n'est pas significatif.

Comme pour toutes enquêtes, les taux de participation sont meilleurs en zone rurale et décroissent avec la taille des villes (cf. graphique 5); le biais géographique, (bien que lié en partie à la répartition généraliste-spécialiste) est l'un des plus préoccupants; un effort particulier devra donc être effectué dans les grandes villes. Le mode d'exercice (cabinet individuel ou de groupe) ne joue pas de rôle significatif.

GRAPHIQUE 5



L'activité est celle donnée par l'entreprise qui nous a fourni le fichier; il s'agit d'une évaluation réalisée par des délégués médicaux travaillant ordinairement pour l'industrie pharmaceutique. Seule, la première classe, avec un taux de participation de 30,9 %, est significativement différente des trois autres (43,3 %).

2. PÉRIODE D'OBSERVATION ET PROFIL HEBDOMADAIRE D'ACTIVITÉ

On connaît l'importance de la saisonnalité sur la mortalité et la morbidité, et, partant, sur la demande de soins médicaux et l'activité des médecins : la représentativité d'une enquête dans ce domaine doit l'être au double plan de la population et de la période de l'année couverte. Ce point a été étudié par ailleurs et n'était pas dans les objectifs de l'enquête expérimentale auprès des médecins (période de relevés : juin-juillet 1980).

On s'intéressera ici au profil hebdomadaire de l'activité des médecins :

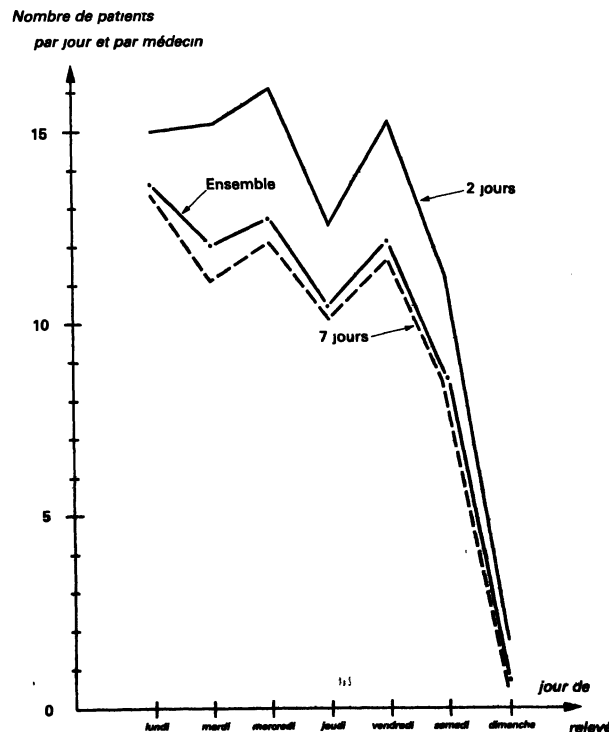
Chaque médecin participant à l'enquête devait remplir les questionnaires séances relatifs à tous les patients vus pendant 2 jours ou 7 jours consécutifs; la date du premier jour de relevé était fixée a priori pour assurer une représentation identique de tous les jours de la semaine. Un certain nombre de médecins n'ont pas strictement suivi ces consignes, des erreurs matérielles ont été commises dans l'envoi des documents et l'échantillon réel diffère légèrement sur ce point de l'échantillon théorique prévu : dans le sous-échantillon 7 jours, la représentativité de tous les jours de la semaine est correctement assurée; au contraire, dans le sous-échantillon de 2 jours, la répartition est très inégale : le mardi représente 27,4 % des jours de relevés et le dimanche seulement 3,7 %.

Un tel biais temporel dans l'observation entraîne une double erreur systématique :

— En niveau, puisque le nombre de patients vus par jour n'est pas uniforme tout au long de la semaine (cf. graphique 6). Le nombre de patients enregistrés par jour et par médecin varie de 13,6 pour

GRAPHIQUE 6

*Profil hebdomadaire de l'activité des médecins*



le lundi à 0,8 pour le dimanche; l'activité la plus forte a lieu en début de semaine : lundi, mardi, mercredi et le vendredi; la plus faible le samedi et le dimanche. Notons la similitude des courbes 2 jours et 7 jours, et le niveau plus élevé de la première dû en partie à la moins grande lassitude en une si courte période des médecins enquêtés, en partie au choix qu'ils ont vraisemblablement fait de leurs 2 jours (consécutifs) de plus grande activité.

— En contenu, puisque les médecins voient vraisemblablement une plus grande porportion d'urgence le samedi et le dimanche, d'enfants le mercredi, ... Ce dernier point n'a pu être vérifié sur l'échantillon trop exigu de l'enquête expérimentale, mais le sera sur les données de l'enquête en vraie grandeur.

### 3. PERTE D'INFORMATION DUE A LA LASSITUDE

Le coût d'une enquête sur le terrain (après sa mise en route et avant la saisie des données) se répartit en un coût fixe (correspondant au contact et à l'obtention de l'adhésion de l'enquêté) et un coût proportionnel (à la quantité d'informations demandée à chaque enquêté). Pour rentabiliser le contact, la tendance générale est d'accroître la charge de travail demandée à chaque participant. Une telle tendance doit être réfreinée car elle entraîne une double réaction des enquêtés :

— La plus visible consiste à réduire le taux de participation : nous avons vu (§ 116) que pour l'échantillon non rémunéré, le taux de participation passait de 43,16 % à 25,99 % lorsque la durée de relevé passait de 2 à 7 jours; de même (§ 117) ce taux passe de 43,09 % à 39,74 %, et 37,03 % lorsque la fiche patient devient plus lourde à remplir.

— La plus difficile à observer, et à redresser, consiste à effectuer le travail demandé de manière de plus en plus négligée; une telle négligence peut prendre de multiples formes, et nous en étudierons deux : le nombre de patients déclarés et les données relevées pour chaque patient.

#### 3.1. *Densité de relevés*

Les relevés des médecins sont, théoriquement, le reflet de leur activité, puisque, pour chaque patient vu, un questionnaire doit être rempli. Les médecins peuvent toutefois se lasser et l'objectif de ce paragraphe est d'analyser la densité de relevés en fonction de la durée d'observation. Comme indicateur de densité, on utilise le nombre de patients pour lesquels une fiche a été remplie par jour.

Les 150 médecins enquêtés théoriquement pendant 7 jours ont enregistré en moyenne 9,7 patients par jour, alors que les médecins enquêtés théoriquement pendant 2 jours ont enregistré en moyenne 14,2 patients. La densité de relevés est donc supérieure de 46 % quand l'observation ne dure que 2 jours (au lieu de 7 jours); cette importante différence est due simultanément au profil hebdomadaire de l'activité des médecins <sup>(1)</sup> (moins de patients le samedi et très peu le dimanche) et à la diminution du nombre de relevés au fur et à mesure que les jours passent.

##### 3.1.1. *Diminution des relevés*

Entre le premier jour de relevé et le septième, le nombre de patients enregistrés, par jour et par médecin, passe de 14,3 à 6,9, soit une diminution <sup>(2)</sup>, de 107,2 %, même si l'observation ne dure que deux jours les enregistrements diminuent de 20,6 % (respectivement 16,4 et 13,6 patients par médecin et par jour) entre le deuxième jour et le premier jour (cf. tableau 9 et graphique 7).

1. Par bonne volonté, certains des médecins deux jours ont vraisemblablement fait pression sur les enquêteurs pour avoir à fournir l'observation sur la partie la plus active de leur semaine : les jours les plus pleins (lundi, mardi, mercredi, vendredi) sont observés 2,4 fois plus souvent que les jours faibles (jeudi, samedi, dimanche).

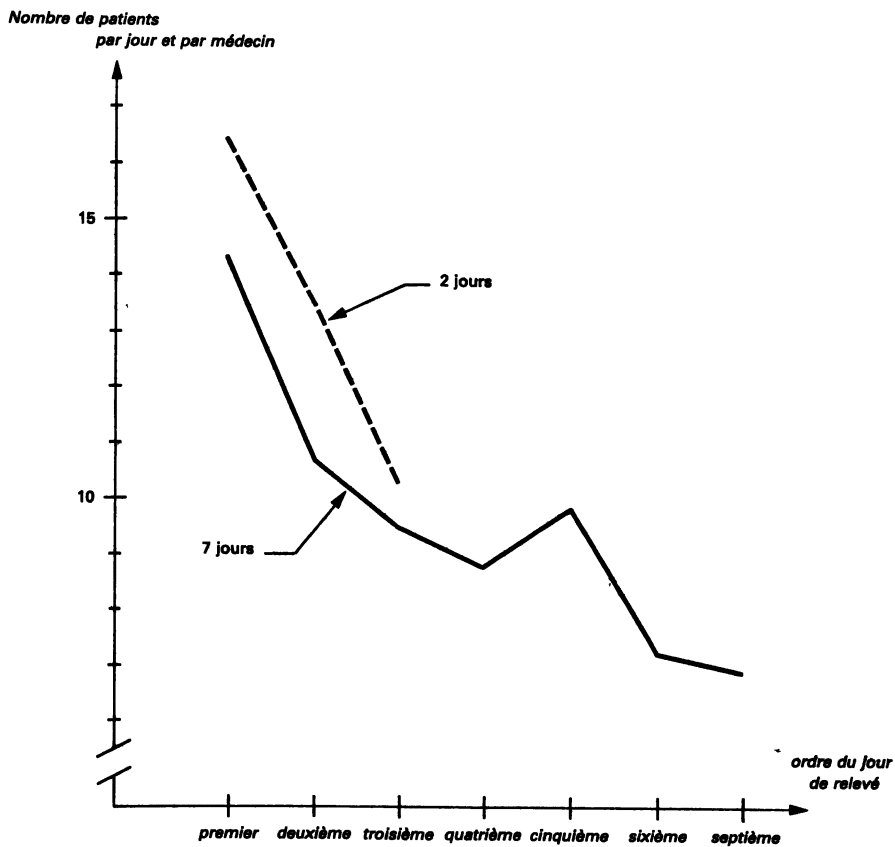
2. La diminution (ou perte) a été calculée, pour chaque jour, par référence au nombre moyen observé le jour considéré.

TABLEAU 9

*Diminution des relevés au cours de la période d'observation  
(en nombre de patients enregistrés par jour et par médecin)*

	2 jours	7 jours	Ensemble	Rapport n <sup>o</sup> jour
				1 <sup>er</sup> jour
1 <sup>er</sup> jour . . . . .	16,4	14,3	15,0	1,00
2 <sup>e</sup> jour . . . . .	13,6	10,7	11,6	0,77
3 <sup>e</sup> jour . . . . .	10,2	9,5	9,5	0,63
4 <sup>e</sup> jour . . . . .		8,8	8,8	0,59
5 <sup>e</sup> jour . . . . .		9,8	9,7	0,65
6 <sup>e</sup> jour . . . . .		7,3	7,4	0,49
7 <sup>e</sup> jour . . . . .		6,9	6,8	0,45
Ensemble . . . . .	14,2	9,7	10,3	0,69

Cette diminution entraîne que le nombre de patients enregistrés le lundi (par exemple) n'est pas le même si le lundi est le premier jour de relevé, le deuxième ou le septième.



GRAPHIQUE 7

*Enregistrement des patients au cours de la période d'observation*

### 3.1.2. Perte due à l'allongement de la période d'observation

Du fait de la représentation inégale des jours de la semaine dans l'échantillon 2 jours, et du profil hebdomadaire de l'activité des médecins, la perte entre 2 jours et 7 jours d'observation ne peut être estimée par la diminution observée du nombre de patients par jour, respectivement 14,2 et 9,7, soit une diminution de 46 %.

Deux estimations de cette perte ont été faites :

1. Si l'échantillon 2 jours avait porté sur les mêmes jours de la semaine que l'échantillon 7 jours, le nombre de patients relevés par jour aurait été de 12,5, soit une perte de 28,9 % environ. Cette estimation a l'inconvénient de donner autant d'importance aux jours peu représentés dans l'échantillon 2 jours (samedi : 13 médecins — dimanche : 6 médecins) qu'aux autres jours.

2. On fait une moyenne des pertes calculées pour chaque jour de la semaine en pondérant par l'activité de chaque jour; la perte ainsi calculée s'élève à 29,8 %.

Quelle que soit la méthode de calcul, la perte due à l'allongement de la période d'observation de 2 jours à 7 jours est de l'ordre de 29 % des patients.

Cette chute considérable de l'enregistrement traduit la lassitude du médecin devant l'effort demandé, mais deux questions essentielles se posent :

- les deux premiers jours d'observation sont-ils le meilleur reflet de l'activité?
- est-ce l'enregistrement de l'activité qui diminue au cours de l'enquête, ou est-ce l'activité elle-même, c'est-à-dire le nombre de patients vus par jour? et comment se fait cette diminution dans les deux cas?

### 3.1.3. Mécanisme de la perte

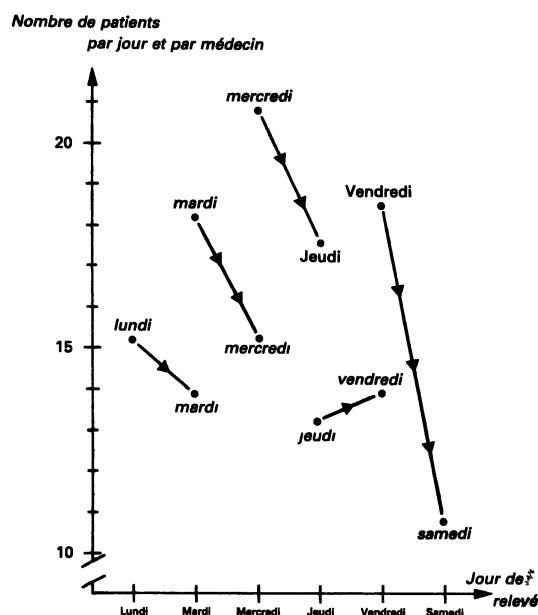
Dans l'échantillon 2 jours et pour les jours de pleine activité, le nombre de patients enregistrés est presque systématiquement inférieur le deuxième jour : par exemple, si le mardi est le premier jour de relevé, 18,2 patients en moyenne sont enregistrés; si le mardi est le deuxième jour (l'enquête ayant commencé un lundi), 13,9 patients seulement sont enregistrés. Le même phénomène se retrouve pour tous les jours d'activité, à l'exception du jeudi premier jour, pour lequel le nombre de patients enregistrés paraît anormalement faible.

Par ailleurs, quatre médecins seulement ont commencé l'enquête un samedi ou un dimanche (cf. tableau 10 et graphique 8).

TABLEAU 10

*Densité de relevés selon les jours et leur rang  
(échantillon 2 jours)  
(entre parenthèses figurent les effectifs de médecins)*

	Nombre de patients par jour et par médecin		
	1 <sup>er</sup> jour début d'enquête	2 <sup>e</sup> jour fin d'enquête	Rapport 1 <sup>er</sup> jour 2 <sup>e</sup> jour
Lundi . . . . .	15,2 (24)		
Mardi . . . . .	18,2 (18)	13,9 (24)	1,31
Mercredi . . . . .	20,8 (6)	15,3 (18)	1,36
Jeudi . . . . .	13,3 (11)	17,5 (6)	0,76
Vendredi . . . . .	18,5 (10)	13,9 (11)	1,33
Samedi . . . . .	12,0 (3)	10,8 (10)	1,11
Dimanche . . . . .	11,0 (1)	0 (3)	—
Lundi . . . . .	—	19 (1)	—
Ensemble . . . . .	16,4 (73)	13,6 (73)	1,21



GRAPHIQUE 8

Nombre de patients relevés par jour selon le jour et son rang (échantillon 2 jours)

La perte moyenne entre le deuxième jour et le premier jour (20,6 %) est perturbée par le profil hebdomadaire d'activité des médecins. Si les deuxièmes jours se répartissaient comme les premiers jours dans la semaine, on aurait observé 15,9 patients par jour (au lieu de 13,6), soit une perte faible entre le premier jour et le deuxième jour, de 3,1 %. En fait, les dimanches et lundis sont insuffisamment représentés (respectivement 3 et 1 médecins) et l'estimation est meilleure si l'on considère les jours du mardi au vendredi; on observe, en moyenne, 17,4 patients le premier jour et, avec la même répartition selon les jours, 15,0 patients le deuxième jour, soit une perte de 16 %.

Pour l'échantillon 7 jours, la perte brute par rapport au premier jour d'observation varie de 33,6 % pour le deuxième jour, à 107,2 % pour le 7<sup>e</sup> jour (cf. tableau 11).

TABEAU 11

Densité de relevés selon les jours et leur rang (échantillon 7 jours) (entre parenthèses figurent les effectifs de médecins)

	Nombre de patients par jour et par médecin							
	1 <sup>er</sup> jour	2 <sup>e</sup> jour	3 <sup>e</sup> jour	4 <sup>e</sup> jour	5 <sup>e</sup> jour	6 <sup>e</sup> jour	7 <sup>e</sup> jour	Ensemble
Lundi . . .	15,4 (54)	4,0 (1)	14,6 (8)	10,6 (19)	12,9 (22)	12,2 (18)	11,5 (22)	13,3 (144)
Mardi . . .	13,9 (22)	11,4 (54)	4,0 (1)	11,0 (8)	10,1 (19)	9,3 (22)	10,2 (18)	11,1 (144)
Mercredi . . .	16,9 (20)	12,6 (22)	12,3 (49)	6,0 (1)	8,1 (9)	9,6 (19)	10,4 (22)	12,1 (14)
Jeudi . . .	12,7 (23)	9,8 (20)	10,3 (22)	10,1 (48)	8,0 (1)	6,3 (8)	9,5 (19)	10,2 (141)
Vendredi . . .	13,0 (20)	11,1 (23)	11,6 (18)	12,0 (22)	11,7 (48)	7,0 (1)	9,4 (8)	11,7 (140)
Samedi . . .	11,6 (8)	11,0 (20)	6,8 (22)	18,3 (18)	9,7 (22)	7,1 (48)	2,0 (1)	8,4 (139)
Dimanche . . .	1,0 (1)	1,4 (8)	0,6 (19)	0,6 (22)	1,4 (18)	0,3 (22)	0,6 (47)	0,6 (137)
<b>Ensemble . .</b>	<b>14,3 (148)</b>	<b>10,7 (148)</b>	<b>9,5 (139)</b>	<b>8,8 (138)</b>	<b>9,8 (138)</b>	<b>7,3 (138)</b>	<b>6,9 (137)</b>	<b>9,7 (986)</b>

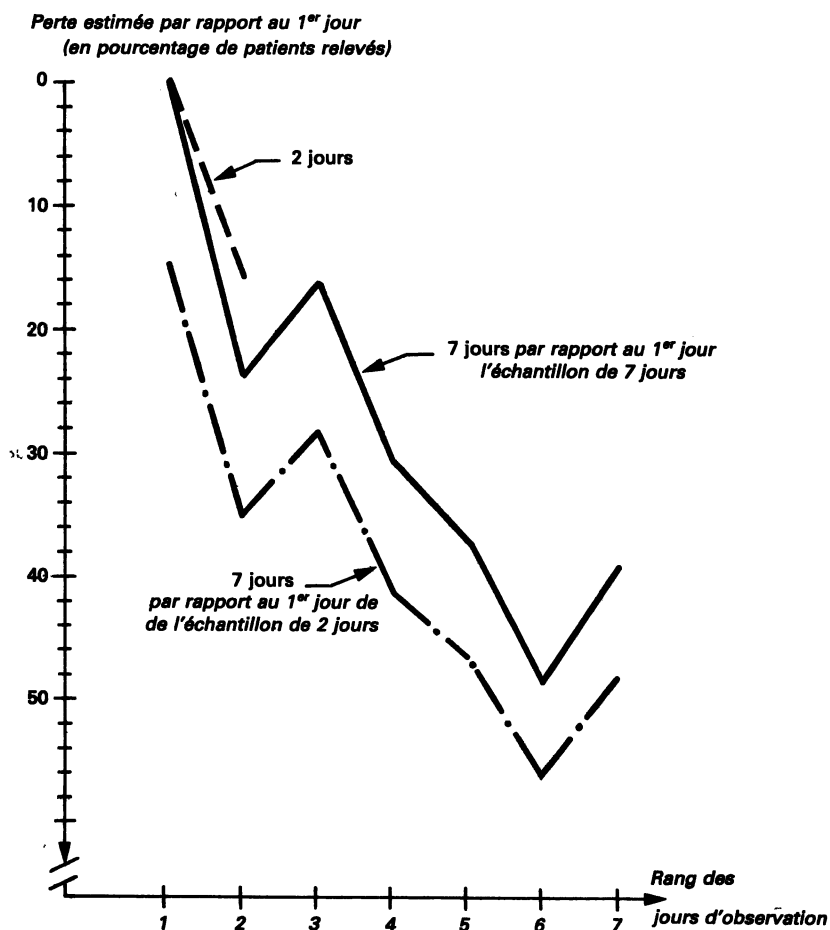


Les perturbations entraînées par le profil hebdomadaire d'activité sont, là aussi, très importantes, et les pertes ont été estimées en gardant la même convention (structure selon les jours de la semaine identique à celle observée pour le premier jour d'enquête, effectif de médecins supérieur à 5). La perte augmente systématiquement avec la durée de l'observation, et atteint environ 40 % au bout de 7 jours (cf. graphique 9).

Le premier jour d'observation n'est pas forcément la meilleure référence pour estimer la perte; ainsi, le nombre de patients relevés le premier jour de l'échantillon 2 jours est 16,4 et celui de l'échantillon 7 jours 14,3 (1), soit une perte de 14,7 %.

GRAPHIQUE 9

*Pourcentage de patients perdus par rapport au premier jour au fur et à mesure du déroulement de l'enquête*



1. Cette comparaison n'est pas biaisée par la représentation différente des jours de la semaine dans les deux cas, car le nombre de patients relevés le premier jour dans l'échantillon 2 jours n'est pas changé (16,4 dans les deux cas), si on applique la structure des jours de la semaine du premier jour de l'échantillon 7 jours.

La perte globale en % du nombre de patients s'élève à :  
 — 8,0 % pour l'échantillon 2 jours,  
 — 38,6 % pour l'échantillon 7 jours.  
 (cf. tableau 12).

TABLÉAU 12

*Estimation (1) de la perte (en % de patients) par rapport au premier jour selon le rang du jour et la durée d'observation*

	Échantillon 2 jours		Échantillon 7 jours		
	Perte brute	Perte estimée	Perte brute par rapport au 1 <sup>er</sup> jour de l'échantillon 7 jours	Perte estimée	Perte estimée par rapport au 1 <sup>er</sup> jour de l'échantillon 2 jours
1 <sup>er</sup> jour . . . . .	0		0		14,7
2 <sup>e</sup> . . . . .	20,6	16,0	33,6	24,1	35,3
3 <sup>e</sup> . . . . .			50,5	16,0	28,3
4 <sup>e</sup> . . . . .			62,5	30,8	41,0
5 <sup>e</sup> . . . . .			45,9	37,4	46,6
6 <sup>e</sup> . . . . .			95,9	48,5	56,1
7 <sup>e</sup> . . . . .			107,2	39,0	48,0
Perte globale . . . . .		8,0		28,0	38,6

1. La perte estimée chaque jour est celle rapportée au nombre moyen observé le jour considéré.

3.2. Contenu des relevés

Le tableau 13 présente trois indicateurs de la quantité d'informations recueillies selon les deux principales variables contrôlées de l'enquête (durée d'enquête et rémunération), et en ventilant selon la nature du médecin (généraliste ou spécialiste).

TABLÉAU 13

*Quantité d'informations recueillies selon la nature du médecin, la durée de l'enquête et la rémunération*

Nature de l'information	Généralistes			Spécialistes		
	2 j	7 J R	7 J R	2 J	7 J R	7 J R
Nombre de diagnostics déclarés par séance . . . . .	1,86	1,71	1,59	1,53	1,56	1,50
Nombre d'ordonnances de pharmacie par séance . . . . .	0,85	0,84	0,79	0,62	0,65	0,68
Nombre de lignes d'ordonnance par ordonnance . . . . .	3,06	2,90	2,99	2,92	2,80	2,47

Il ne semble pas que les modalités de l'enquête modifient le comportement du médecin vis-à-vis de son patient : le nombre d'ordonnances par séance, et le nombre de lignes par ordonnance ne varient de manière systématique, ni avec la durée de l'enquête, ni avec la rémunération du médecin; pour les

spécialistes, cependant, le nombre de lignes par ordonnance est faible pour l'échantillon 7 J R (15,4 % de moins que 2 J et 11,8 % de moins que 7 J R) : En dehors d'une erreur aléatoire, un tel écart peut être dû aux difficultés de manipulation des questionnaires. Un tel résultat est encore plus net pour le travail proprement dit, entraîné par la participation à l'enquête pour le généraliste comme pour le spécialiste, le nombre de diagnostics est légèrement plus faible pour l'échantillon 7 J R.

Au total, comme pour les taux de participation et pour le nombre de patients vus chaque jour (densité de relevés), le contenu de l'information recueillie pour chaque patient est amélioré par le versement d'une rémunération et se dégrade lorsque la durée d'observation se prolonge.

### CONCLUSION

Remarquons tout d'abord que l'observation recueillie au cours de l'enquête auprès des médecins de 1979 se situe à un double niveau, puisqu'elle porte simultanément sur un échantillon de médecins (que nous appelons médecins-enquêtés) et sur un échantillon de leurs patients (plus ou moins important selon la durée de relevé); les médecins-enquêtés jouent, de ce fait, un double rôle, d'enquêtés, puisqu'ils sont astreints à remplir un questionnaire médecin, et d'enquêteur, puisqu'ils nous fournissent des fiches patients pour tous les clients vus durant la période d'observation : Dans les enquêtes habituelles, où l'enquêteur et l'enquêté sont disjoints, la mesure du biais doit être analysée séparément aux deux niveaux.

Le choix d'un protocole d'enquête est le résultat d'un compromis tenant compte d'un grand nombre de contraintes, parmi lesquelles la minimisation des biais n'est pas la moindre, et les modalités de chaque enquête dépendront donc de la nature des recherches qu'on envisage d'effectuer et de l'importance qu'on accorde aux différents types d'erreurs; quelles que soient ces modalités, l'importance de bien convaincre toutes les personnes appartenant à l'échantillon choisi, de l'intérêt de l'enquête, apparaît comme un résultat général; pour cela, différents moyens pourront être mis en œuvre : campagne de presse en direction de l'ensemble de la population, prospectus envoyés aux enquêtés, rémunération, retour d'information, allègement autant que possible du travail ou du dérangement qui leur est imposé, ... Enfin, une information de bonne qualité coûte cher et on aurait souvent intérêt à réduire la taille de l'échantillon, et/ou la quantité d'informations obtenues auprès de chaque enquêté, si ces mesures permettent d'améliorer, à budget constant, la fiabilité des données.

## BIBLIOGRAPHIE

- [1] C. GUILLOT, An. et Ar. MIZRAHI. — Étude critique de méthodes d'enquête (une enquête pilote sur la consommation médicale). *Consommation*, n° 1, janvier-mars 1968.  
C. GUILLOT, An. et Ar. MIZRAHI. — Enquête pilote 1965-1966 sur les soins médicaux. Fascicule 1 : résultats méthodologiques. C.R.E.D.O.C., mars 1968 (rapport à tirage limité).
- [2] An. et Ar. MIZRAHI. — Précision et saisonnalité. C.R.E.D.O.C., mars 1969 (rapport à tirage limité).  
An. et Ar. MIZRAHI. — Durée d'observation et précision dans les enquêtes de consommation. *Consommation*, n° 4, 1969.
- [3] An. et Ar. MIZRAHI. — L'enquête sur la santé des Canadiens. Aspects techniques de l'entrevue et appariement avec les fichiers d'assurance maladie et hospitalisation. Rapport de mission, avril 1975, Ottawa (texte français, 55 pages). Rapport interne.
- [4] An. et Ar. MIZRAHI. — Méthode d'enquête : l'observation des hospitalisations auprès des ménages et dans les dossiers de la Sécurité sociale. C.R.E.D.O.C., avril 1976 (58 p.).
- [5] An. et Ar. MIZRAHI. — Méthode d'enquête : fiabilité de la mémoire des enquêtés (hospitalisation sur une année). C.R.E.D.O.C., mai 1976 (18 pages). Rapport interne.
- [6] Th. LECOMTE, An. et Ar. MIZRAHI. — La morbidité hospitalière : méthode d'enquête. C.R.E.D.O.C., novembre 1976 (40 pages).
- [7] An. et Ar. MIZRAHI, A. ZOUAOU. — Projet de sondage dans les fichiers de Sécurité sociale. C.R.E.D.O.C., février 1977 (39 pages).  
An. et Ar. MIZRAHI. — Méthode de sondage — Enquête permanente dans les dossiers de Sécurité sociale (Panel) C.R.E.D.O.C., janvier 1978 (50 pages).  
An. et Ar. MIZRAHI - A. ZOUAOU. — Présentation d'un sondage représentatif dans les fichiers de Sécurité sociale. *Leçons d'économie médicale (Collection de Médecine légale et de Toxicologie médicale)*, n° 116, 3<sup>e</sup> trimestre 1980 (12 pages).
- [8] Arié MIZRAHI. — Approches de la mesure de la morbidité de la population à partir d'enquêtes auprès des ménages. Texte présenté à la Conférence-séminaire sur les systèmes d'enquêtes nationales de santé dans la Communauté Économique Européenne (Bruxelles, octobre 1975) repris en anglais dans la Publication EUR 5747e/1977 de la C.E.E.
- [9] Th. LECOMTE, An et Ar. MIZRAHI. — Projet d'enquête sur la thérapeutique médicale, C.R.E.D.O.C., novembre 1979 (52 pages).
- [10] Th. LECOMTE, An et Ar. MIZRAHI, D. PILLIARD. — Enquête sur les hospitalisés : socio-démographie, pathologie, consommation. *Méthodologie*. C.R.E.D.O.C., juin 1980 (200 pages).
- [11] Th. LECOMTE, An. et Ar. MIZRAHI. — Présentation de l'enquête nationale 1980-81 sur les consommations médicales. C.R.E.D.O.C., juillet 1980 (30 pages).  
Th. LECOMTE. — Enquête sur la santé et les soins médicaux 1980-81 : relevé et codification des données de morbidité. C.R.E.D.O.C., novembre 1981 (106 pages).
- [12] An. et Ar. MIZRAHI. — Les enquêtes nationales auprès des ménages sur la consommation médicale en France. Journées d'information sur le Programme National Suisse de Recherche — Efficience en santé publique, Bâle, 6-7 novembre 1980. C.R.E.D.O.C., novembre 1980 (23 pages).  
An. et Ar. MIZRAHI. — Système d'observation micro-économique de la consommation médicale en liaison avec la morbidité en France. Communication « Secondo Incontro Internazionale di studio Sulla Riforma Sanitaria Consiglio Nazionale Dell'Economica e del Lavarò », Rome, 2-4 juin 1981. C.R.E.D.O.C., avril 1981 (20 pages).  
An. et Ar. MIZRAHI. — Système d'observation micro-économique de la consommation médicale en liaison avec la morbidité en France. Colloque « Franco-Sociétique », Moscou, 18-24 mai 1981. C.R.E.D.O.C., mai 1981 (14 pages).
- [13] Ph. LE FUR, An. et Ar. MIZRAHI. — Méthode d'enquête : morbidité et thérapeutique médicale. C.R.E.D.O.C., octobre 1981 (231 pages).
- [14] An. et Ar. MIZRAHI. — La consommation médicale : micro-économie, P.U.F., collection l'Économiste, 1<sup>er</sup> trimestre 1982 (267 pages).