

# À propos de l'estimation de la prévalence et de l'incidence des maladies héréditaires

Josué Feingold

*Les généticiens utilisent fréquemment le terme de fréquence pour mesurer dans une population l'importance d'un gène, d'un génotype, d'un phénotype normal ou pathologique. Le terme de fréquence est correct s'il concerne un polymorphisme neutre ou quasi neutre vis-à-vis de la sélection naturelle, tels les groupes sanguins ABO ou Rh, les polymorphismes de l'ADN, etc. Dans ces systèmes génétiques les fréquences des gènes, des génotypes et des phénotypes ne varient pratiquement pas avec l'âge des individus. En fait, il s'agit, dans ces cas, d'une mesure de la prévalence du gène, du génotype ou du phénotype. Le terme de fréquence est, en revanche, très imprécis lorsque l'on considère les maladies héréditaires, c'est-à-dire celles qui se*

*transmettent selon les lois de Mendel. Il faut dans ce domaine mesurer la morbidité du caractère pathologique en utilisant les trois indices définis par les épidémiologistes : la prévalence, l'incidence et l'incidence cumulée. Dans ce dossier, nous définirons ces trois indices de morbidité, nous montrerons à l'aide d'exemples simples comment ils peuvent être calculés dans les maladies héréditaires. Enfin, nous ferons quelques commentaires concernant ces trois paramètres. Nous n'envisagerons pas les problèmes posés, d'une part, par les maladies par aberrations chromosomiques et, d'autre part, par les maladies multifactorielles c'est-à-dire celles qui ont deux composantes l'une génétique, l'autre liée à l'environnement.*

## Définition des indices de morbidité

### • La prévalence

La prévalence [1-3] est un indice de morbidité qui concerne les malades présents dans une population à un moment ou durant une période donnée. Deux types de prévalence peuvent être estimés : la prévalence instantanée et la prévalence au cours d'une période de temps. La prévalence instantanée  $P_i$  représente la proportion de personnes malades parmi une population à un instant donné. Si  $M$  est le nombre de malades et  $N$  le nombre total de sujets de la population,  $P_i = M/N$ . La prévalence  $P_p$  au cours d'une période de temps  $p$  est le rapport entre le nombre de personnes malades à un moment quelconque

de la période et l'effectif moyen de la population au cours de la période, celle-ci est habituellement d'un an dans les maladies chroniques.

La mesure de la prévalence peut concerner l'ensemble de la population ou une fraction bien définie : sujets de sexe masculin ou féminin, sujets à la naissance, ou appartenant à une tranche d'âge donnée. La mesure de la prévalence à la naissance est également appelée incidence à la naissance par de nombreux auteurs. On peut noter que la mesure de la fréquence des polymorphismes neutres est en fait une mesure de leur prévalence.

### • L'incidence

Le taux annuel d'incidence dans une population est le plus utilisé pour les maladies chroniques [1-3]. Il se défi-

nit par le rapport : nombre de nouveaux cas recensés pendant une période donnée / nombre total de personnes-années. Si la population est stable et la maladie étudiée rare, le dénominateur du rapport est le produit du nombre d'années d'observations par l'effectif de la population. Le taux d'incidence  $T_i$  est dit spécifique s'il concerne une fraction définie de la population : sujets de sexe masculin ou féminin, appartenant à une tranche d'âge donnée, ou ayant un génotype donné.

### • L'incidence cumulée

Les taux d'incidence spécifique par âge permettent de calculer l'incidence cumulée d'une maladie de la naissance à un âge donné  $t$  ou entre deux âges  $t_1$  et  $t_2$  [1-3]. Si la maladie est rare l'incidence cumulée est la

somme des incidences spécifiques de la période considérée. L'estimation d'incidence cumulée a deux applications possibles dans les maladies héréditaires :

- Si l'incidence cumulée concerne, d'une part, l'ensemble de la population (ou un échantillon représentatif) et, d'autre part, la vie entière des sujets, elle estime la prévalence à la naissance des sujets porteurs du génotype pathologique, à la condition que la pénétrance du caractère soit complète. Si cette dernière est incomplète l'estimation doit être corrigée. Rappelons que la pénétrance est la probabilité pour un sujet ayant un génotype délétère de faire la maladie.

- L'incidence spécifique par âge et l'incidence cumulée, si elles concernent uniquement les sujets porteurs du génotype délétère, permettent d'estimer la pénétrance. L'incidence spécifique, dans ce cas, concerne la distribution de l'âge du début de la maladie, le second indice est la pénétrance en fonction de l'âge et en particulier la pénétrance finale du caractère. Par exemple, dans la maladie de

Huntington dont l'âge du début est variable, la pénétrance est de 1 vers l'âge de soixante-dix ans, mais de zéro à la naissance et d'environ 0,50 à quarante ans [4].

Le *Tableau I* emprunté à Khoury *et al.* [1] résume les caractéristiques des trois indices.

### Quelques exemples de calcul des indices de morbidité

Ces estimations peuvent être le résultat d'enquêtes ou d'un recensement systématique des cas.

#### • Prévalence à la naissance

Elle concerne selon les cas la prévalence de la maladie ou celle du génotype à l'origine de la maladie. Par exemple, la phénylcétonurie est une maladie récessive autosomique qui fait l'objet d'un dépistage néonatal. Le dépistage concerne le trouble métabolique (phénylalaninémie augmentée) et non le phénotype pathologique qui apparaît plus tardivement. Du 1/1/1978 au

31/12/1991, 10 724 488 nouveau-nés en France ont été testés, 637 se sont révélés être atteints d'où une prévalence à la naissance de 1/16616 [5].

La prévalence à la naissance de la mucoviscidose dans la population du Nord-Finistère a été estimée rétrospectivement parmi les 60 766 enfants nés entre le 1/1/1965 et le 31/12/1972. Trente-quatre cas ont été recensés, les sources d'information ont été multiples. La prévalence à la naissance des sujets porteurs du génotype pathologique est donc de 34/60766 soit  $5,6 \cdot 10^{-4}$  [6].

#### • Prévalence d'une maladie héréditaire dans la population

Nous prendrons comme exemple l'estimation de la prévalence de la maladie de Huntington dans la population du sud du Pays de Galles. Il s'agit, rappelons-le, d'une maladie dominante autosomique. La prévalence instantanée a été estimée par Walker *et al.* [7]. Le jour choisi a été le 25-26 avril 1971, date à laquelle la population a été recensée. Le

Tableau I

### CARACTÉRISTIQUES DES DIFFÉRENTES MESURES DE LA FRÉQUENCE D'UNE MALADIE

	Incidence	Incidence cumulée	Prévalence
<b>Numérateur</b>	Nombre de nouveaux cas	Nombre de nouveaux cas durant une période donnée (la vie entière par exemple)	Nombre de cas au moment de l'étude
<b>Dénominateur</b>	Somme des personnes-temps	Taille de la population au début de la période	Taille de la population au même moment
<b>Dimension</b>	1/temps $\rightarrow t^{-1}$	Aucune	Aucune
<b>Autres dénominations</b>	Densité d'incidence, force	Risque de maladie, probabilité de maladie	Fréquence des sujets malades ou d'un caractère
<b>Exemples</b>		Risque durant la vie Pénétrance	Fréquence génique ou génotypique d'un marqueur Fréquence à la naissance d'une maladie
<b>Commentaires</b>	Paramètre important en épidémiologie	Paramètre important en épidémiologie, génétique	Paramètre utile en santé publique et en génétique des populations

D'après [6].

nombre de malades était ce jour de 131, la population étant formée de 1720901 sujets, la prévalence de la maladie était donc de  $7,61 \cdot 10^{-5}$ .

Il faut noter que le recueil des informations médicales a été rétrospectif, a duré cinq ans (1973-1978), et que les sources d'informations ont été multiples. A partir de cette prévalence on peut calculer la prévalence des hétérozygotes (malades et non malades). Différentes méthodes ont été décrites pour réaliser ce calcul; elles prennent toutes en compte la distribution de l'âge du début de la maladie [8].

La prévalence des hétérozygotes est, dans le cas de la maladie de Huntington, d'environ 2,5 fois celle de la maladie. Cette estimation mesure la prévalence à la naissance des sujets ayant le génotype pathologique. Elle est d'environ  $19 \cdot 10^{-5}$  au sud du Pays de Galles.

Ce type d'enquête a été critiqué car le recensement des malades n'est pas exhaustif. La méthode de « captures, recaptures » qui prend en compte les malades recensés par des sources différentes d'information a été préconisée pour corriger ce biais [8]. Elle dérive de la technique utilisée pour estimer le nombre d'animaux sauvages dans une région donnée d'où son nom. Elle n'est cependant pas exempte de critique [9].

#### • Incidence d'une maladie héréditaire

Cet indice est peu utilisé en génétique car, contrairement à la situation en épidémiologie, il ne mesure pas l'impact d'un facteur étiologique. En effet, il concerne le passage du phénotype normal au phénotype pathologique chez les sujets porteurs du génotype délétère. Cependant, il peut être intéressant si l'on veut étudier les facteurs génétiques ou de milieu qui ont un rôle dans l'âge d'apparition du phénotype pathologique. Sa mesure est importante si l'on veut estimer l'incidence cumulée sur une vie entière.

A titre d'exemple, nous décrirons l'enquête qui a permis d'estimer l'incidence par âge de l'adrénoleucodystrophie, une maladie récessive liée au chromosome X [10]. Nous

avons considéré l'ensemble des garçons nés en France entre le 1<sup>er</sup> janvier 1956 et le 31 décembre 1986. Dans cette cohorte, 14 enfants ont débuté la maladie entre 7 et 8 ans; la population étant de 9265300, l'incidence estimée est de  $1,446 \cdot 10^{-5}$ .

#### • Incidence cumulée d'une maladie héréditaire

Elle permet, si l'on considère la vie entière, d'estimer comme nous l'avons dit la prévalence à la naissance du génotype délétère ou la pénétrance du caractère. C'est ainsi qu'en sommant les incidences par âge, nous avons estimé la prévalence à la naissance du génotype délétère dans l'étude de l'adrénoleucodystrophie [10]; elle est d'environ 1/100000 nouveau-nés de sexe masculin (chiffre sous-estimé).

La pénétrance de la maladie de Huntington a été estimée en étudiant la distribution de l'âge du début clinique de la maladie. Elle est de 1 vers l'âge de 70 ans [4]. Pour une des formes du cancer du sein héréditaire une estimation indirecte a permis de montrer que la pénétrance est de 0,80 à l'âge de 70 ans [11].

### Remarques et conclusions

La prévalence par âge est un paramètre important en génétique des populations car elle permet d'étudier les effets de la sélection naturelle (Tableau II). En revanche, en épidémiologie analytique, c'est l'incidence par âge qui est le paramètre utile pour l'étude d'un facteur étiologique. Cette différence entre les deux disciplines dans le domaine de la recherche est due au fait que la relation « facteur étiologique-maladie » n'est pas de même nature selon que l'on considère une maladie non héréditaire (domaine de l'épidémiologie) et une maladie héréditaire (domaine de la génétique). En effet, dans le premier cas, le facteur étiologique (le tabac par exemple) et la maladie (le cancer du poumon) sont des phénomènes de nature totalement différentes: en revanche, dans le second cas, le génotype (cause de la maladie héréditaire) et le phénotype sont deux aspects du même fait, même si la liaison n'est pas totale en cas de pénétrance incomplète.

Pour conclure, rappelons que le terme de « fréquence d'une maladie héréditaire » ne devrait plus être employé. Il

Tableau II

PRÉVALENCE EN FONCTION DE L'ÂGE DES SUJETS ET SELON LEUR TYPE D'HÉMOGLOBINE DANS LA POPULATION DE GARKI (NIGERIA) [12]

Type d'hémoglobine (électrophorèse)	Nouveau-nés (534)	Population générale (2742)
AA	0,738	0,7020
AS	0,236	0,2892
AC	0,006	0,0069
SC	0,000	0,0015
CC	0,000	0,0000
SS	0,020	0,0004
<b>Fréquence génique</b>		
A	0,859	0,850
S	0,138	0,146
C	0,003	0,004

Les sujets SS ont une drépanocytose. On peut noter que les fréquences géniques varient peu avec l'âge mais que, en revanche, les prévalences des génotypes sont différentes. Ces chiffres montrent que la sélection naturelle est forte contre les sujets drépanocytaires et est en faveur des sujets hétérozygotes AS.

faut utiliser les termes de prévalence ou d'incidence et préciser si le paramètre concerne le génotype délétère ou le phénotype pathologique ■

## RÉFÉRENCES

1. Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of genetic epidemiology*. New York, Oxford: Oxford University Press, 1993.
2. Leclerc A, Papoz L, Breart G, Lellouch J. *Dictionnaire d'épidémiologie*. Paris: Editions Frison-Roche, 1990.
3. Rumeau-Rouquette C, Blondel B, Kaminski M, Breart G. *Epidémiologie: méthodes et pratique*. Paris: Flammarion Médecine Sciences, 1993.
4. Harper P. *Huntington's disease*. London: WB Saunders Company Ltd, 1991.
5. Briard ML, Farriaux JP. Le dépistage néonatal de la phénylcétonurie (*Association pour le dépistage et la prévention des handicaps de l'enfant*). *La Dépêche*, novembre 1992; 18: 7-11.
6. Feingold J, Hennequet A, Jehanne M, Feigelson J, Toudic L, Qiniou O, Briard ML. Fréquence de la fibrose kystique du pancréas en France. *Ann Genet* 1974; 17: 257-9.
7. Walker DA, Harper PS, Wells CEC, Tyler A, Davies K, Newcombe RG. Huntington's chorea in South Wales. *Clin Genet* 1981; 19: 213-21.
8. Hook EB, Regal RR. The value of capture-recapture methods even for apparent exhaustive surveys. The need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. *Am J Epidemiol* 1992; 135: 1060-7.
9. Papoz L, Balkau B, Lellouch J. Case counting in epidemiology: limitations of methods based on multiple data sources. *Int J Epidemiol* 1996; 25: 474.
10. Sereni C, Paturneau-Jouas M, Aubourg P, Baumann N, Feingold J. Adrenoleukodystrophy in France: an epidemiological study. *Neuroepidemiology* 1993; 12: 229-33.
11. Easton DF, Bishop DT, Ford D, Crookford GP, and the Breast Cancer Linkage Consortium. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. *Am J Hum Genet* 1993; 52: 678-701.
12. Molineaux L, Gramiccia G. *The Garki Project. Research of the epidemiology and control of malaria in the Sudan Savanna of West Africa*. Geneva: World Health Organisation, 1980.

### Josué Feingold

Directeur de recherche émérite à l'Inserm. Unité de recherches d'épidémiologie génétique, Inserm U. 155, Université Paris VII, 2, place Jussieu, 75005 Paris, France.

### TIRÉS À PART

J. Feingold.



**Le Centre International de Recherche sur le Cancer (IARC/WHO)**  
**The National Institute of Environmental Health Sciences, USA (NIEHS)**  
 organisent un *Symposium International*

**CELL ADHESION AND COMMUNICATION  
 IN GROWTH CONTROL AND CANCER**  
 Lyon, France  
 19-21 janvier 1999

#### Orateurs et principaux thèmes :

- The cell adhesion apparatus (W. Franke)*
- The molecular changes in e-cadherin and wnt pathways in cancers (G. Christofori, S. Hirohashi, P. Polakis\*, P. Guilford, F. van Roy)*
- Cell adhesion and signal transduction in growth control (J.C. Barrett, W. Birchmeier, H. Clevers, T. Noda)*
- Gap junctions (N.B. Gilula)*
- Connexins in growth control and cancer (M. Mesnil, K. Willecke, H. Yamasaki)*
- Connexins and growth control mechanisms (V. Krutovskikh, A.F. Lau, W.H. Moolenaar, C. Naus, B. Rousset)*
- Integrin-related signalling pathways in growth control and cancer (F. Giancotti, E. Ruoslahti, J.P. Thiery)*

Renseignements et inscriptions IARC WebSite : [www.iarc.fr](http://www.iarc.fr)  
 e-mail to : [dechaux@iarc.fr](mailto:dechaux@iarc.fr)  
 Fax : [33] (0) 472 73 84 42 / Tél. : [33] (0) 472 73 85 39