



# Quels risques, pour quelles femmes ?

## Risque populationnel de cancer du sein

### *Breast cancer risk in population*

**Mots-clés :** Épidémiologie – Incidence – Risque – Cancer du sein.

**Keywords:** *Epidemiology – Incidence – Risk – Breast cancer.*

**P. Roy\***

L'estimation du risque populationnel de cancer du sein chez la femme nécessite de définir préalablement la notion de risque. Évaluer un risque, c'est estimer la probabilité de survenue d'un événement. Rappelons simplement qu'il s'agit d'associer à chaque événement sa probabilité, c'est-à-dire un réel positif appartenant à l'intervalle  $[0;1]$ . S'agissant d'une probabilité, il n'y a pas d'unité : la valeur 0 caractérise l'événement impossible, la valeur 1 l'événement certain. Plus l'événement est probable, plus la valeur de la probabilité est élevée. Deux notions se distinguent immédiatement : la première répond à la question de la proportion de cancers du sein dans la population à une date donnée – c'est la prévalence – ; la seconde répond à la question du risque de survenue de cancer dans un groupe d'individus initialement indemnes observés durant un certain temps. Cette notion nous amène à étudier comment se distribue l'âge de survenue des cancers et nécessite l'introduction du taux d'incidence [1].

### La prévalence

---

La prévalence est estimée, dans une population bien définie, en rapportant lors d'études transversales l'effectif des cas présents (prévalents) à l'effectif de la population concernée. Cet indicateur est intéressant pour connaître l'état de santé d'une population à un moment donné et planifier l'offre de soins. La prévalence dépend de l'incidence de la maladie, de la survie des patients et de l'évolution démographique

---

\* Service de biostatistique santé UCBL, équipe de l'UMR CNRS 5558, CHU Lyon-Sud.

de la population d'étude. Tandis que la prévalence totale prend en compte l'ensemble des patients vivants à un moment donné, les prévalences partielle et actuelle sont restreintes aux seuls patients nécessitant des soins [2]. La prévalence n'est pas un taux ; il s'agit d'un indicateur relativement peu étudié à ce jour. L'activité de recherche des registres de population a concerné initialement l'étude de l'incidence de cancer, puis celle de la survie des patients. La modélisation de la prévalence est certainement appelée à se développer dans les années à venir.

## **Le taux d'incidence**

---

L'incidence est l'indicateur de référence des études étiologiques qui identifient les facteurs de risque de survenue des cancers. L'étude du taux d'incidence est particulièrement intéressante pour estimer le risque populationnel de cancer du sein. La nature temporelle du phénomène étudié est essentielle et concerne en premier lieu l'âge de survenue du cancer. La variable aléatoire  $T$  représente l'âge au diagnostic. Sur un intervalle de temps  $dt$  très court, le taux d'incidence instantané est le rapport de la probabilité de survenue du cancer sur cet intervalle, c'est-à-dire entre les âges  $t$  et  $t + dt$  ( $t \leq T < t + d$ ) dans une population préalablement indemne ( $T > t$ ), divisée par  $dt$  :

$$\ddot{e}(t) = \lim_{dt \rightarrow 0} \frac{\Pr(t \leq T + dt < T > t)}{dt}$$

Cette formulation mathématique du taux d'incidence instantané est similaire à celle du taux de mortalité instantané lors de l'étude de la mortalité globale, ce qui n'est pas surprenant : de la même façon que la force de mortalité retire les individus de la cohorte au moment de leur décès, la force d'incidence retire au moment où ils surviennent les cas de la cohorte des sujets indemnes. En pratique, cela permet d'appliquer des résultats théoriques issus de la modélisation de la mortalité à celle de l'incidence. Le taux d'incidence a donc une dimension dynamique. Son unité est l'inverse de l'unité de temps (temps<sup>-1</sup>). Au niveau populationnel, le taux d'incidence est estimé en rapportant le nombre de nouveaux cas (incidents) à la somme des durées d'observation des individus de la cohorte. Le dénominateur est connu sous le nom de "personnes-années d'observation" (années d'observation accumulées par l'ensemble des personnes étudiées). Le taux d'incidence mesure l'intensité avec laquelle une maladie frappe une population.

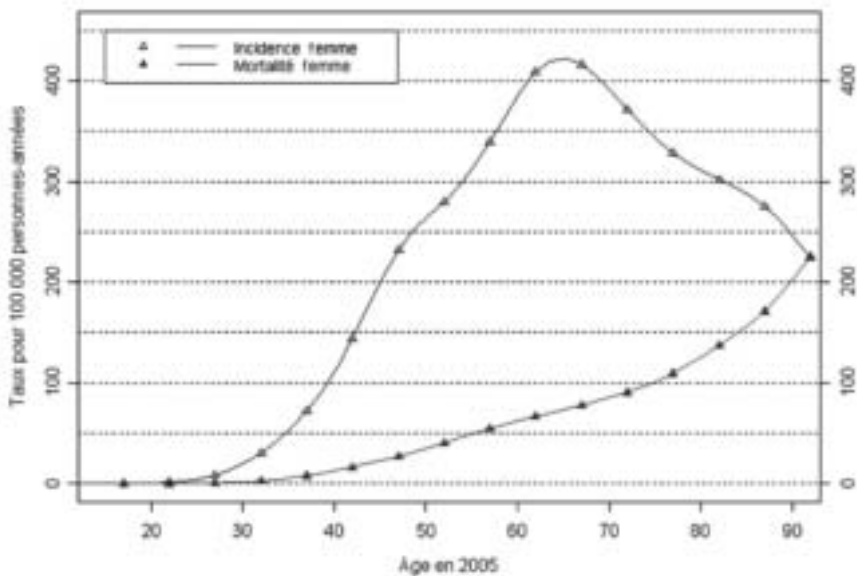
## **Taux d'incidence spécifique par âge / Taux standardisé**

---

Sur une population donnée, le taux brut est le rapport du nombre de cas incidents sur le nombre de personnes-années d'observation. L'influence forte de l'âge justifie l'utilisation du taux d'incidence spécifique par âge. L'estimation de l'incidence du

cancer du sein effectuée à partir des données des registres français du cancer (réseau FRANCIM en partenariat avec le service de biostatistique des hospices civils de Lyon et l'InVS) permet de retrouver une forte influence de l'âge [1]. Alors que les taux d'incidence spécifique par âge montrent un pic de l'ordre de 400 pour 100 000 personnes-années entre 60 et 69 ans pour décroître ensuite, les taux de mortalité par cancer du sein augmentent régulièrement avec l'âge (*figure 1*). L'âge (A), la période (P) et la cohorte de naissance (C) étant reliés par la relation  $A = P - C$ , cette représentation transversale de l'effet de l'âge doit être interprétée en gardant à l'esprit qu'à chaque valeur du taux d'incidence correspond une cohorte de naissance différente.

**FIGURE 1.** Incidence et mortalité par âge en 2005 (d'après [1]).



Les taux d'incidence du cancer du sein variant fortement avec l'âge, des différences de structures d'âge des populations comparées pourraient conduire à observer de fortes différences des taux bruts, alors même que les taux d'incidence spécifiques par classe d'âge seraient voisins, voire identiques. La méthode de standardisation directe consiste à appliquer les taux d'incidence spécifiques par tranches d'âge observés dans la population étudiée aux effectifs correspondants d'une population théorique de référence, à calculer le nombre de cas qui seraient produits dans chaque classe d'âge de cette population théorique, à sommer ceux-ci, puis à diviser le nombre total de cas ainsi calculés par le total des personnes-années de la population de référence. On obtient ainsi le taux standardisé sur l'âge, qui ne dépend plus de la structure d'âge

de la population. L'estimation française du taux brut d'incidence annuelle de cancer du sein pour l'année 2005 est de 159,2 pour 100 000 personnes-années, les valeurs des taux standardisés sur les populations européennes et mondiales étant respectivement de 136,9 et 101,5 pour 100 000 personnes-années. Les valeurs correspondantes des taux de mortalité par cancer du sein sont respectivement de 35,8, 25,5 et 17,7 pour 100 000 personnes-années [1]. L'utilisation des taux standardisés permet d'effectuer des comparaisons entre différents registres au niveau international.

## Incidence cumulée / Risque cumulé

Estimer le risque de cancer n'a de sens que si l'on prend en compte la nature temporelle du phénomène étudié. Le premier facteur temporel à étudier est évidemment l'âge. Les taux de mortalité et d'incidence ne sont pas des probabilités mais sont utilisés pour calculer les probabilités de décès et de survenue d'un cancer entre la naissance et l'âge  $t$ . Le risque cumulé de développer un cancer entre la naissance et l'âge  $t$  est la valeur prise par la fonction de répartition au temps  $t$ :  $R(t)$ . Ce risque cumulé de cancer à l'âge  $t$  est alors estimé à partir des différentes valeurs des taux d'incidence spécifique de l'âge  $\ddot{e}(t)$  en utilisant les relations suivantes :

$$\begin{aligned} \ddot{e}(t) &= \frac{R'(t)}{1 - R(t)} \\ \ln(1 - R(t)) - \int_0^t \ddot{e}(u) du &= -\ddot{E}(t) \\ R(t) &= 1 - e^{-\ddot{E}(t)} \end{aligned}$$

où  $R'(t)$  est la densité de probabilité (la dérivée de la fonction de répartition), et  $\ddot{E}(t)$  le taux d'incidence cumulé à l'âge  $t$ .

## Évolution temporelle de l'incidence

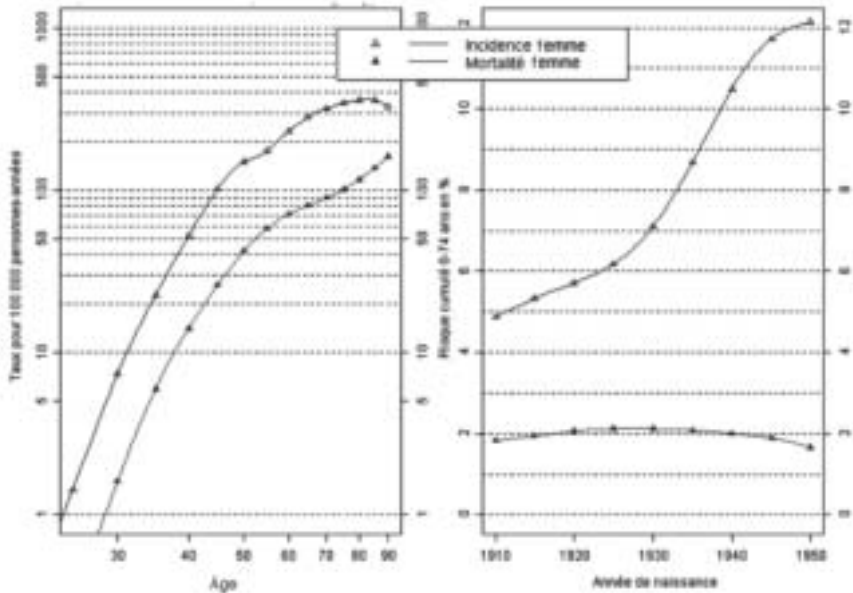
L'objectif de cette modélisation est d'analyser l'évolution du risque de cancer au cours du temps. Les données fournies par les registres français du cancer ont été analysées à l'aide d'un modèle âge/période/cohorte, les effets de l'âge et de la cohorte étant pris en compte à l'aide de fonctions de lissage de type spline, et l'effet période à l'aide d'une fonction polynome de degré 2 [1].

L'évolution du risque cumulé par cohorte de naissance est la plus intéressante, puisque les femmes d'une même génération partagent un certain nombre de facteurs de risque (facteurs environnementaux, habitudes de vie...). La **figure 2a** fournit une représentation longitudinale de l'effet de l'âge pour une même cohorte de naissance. Le risque cumulé de développer un cancer du sein entre 0 et 74 ans pour cette cohorte est estimé à partir de ces valeurs. Ce risque augmente régulièrement avec la cohorte de naissance, pour atteindre 12 % pour les femmes nées en 1950 (**figure 2b**). Cette

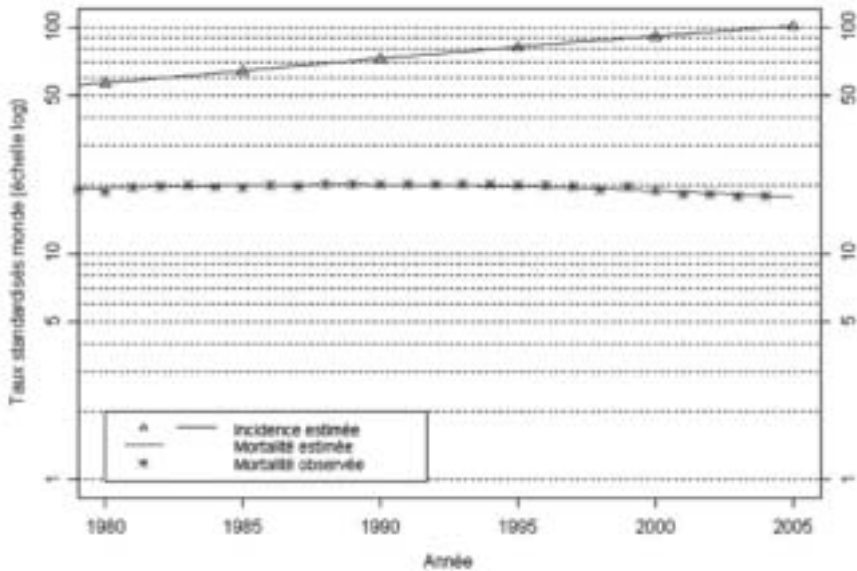
augmentation est particulièrement accentuée pour la génération des femmes nées entre 1920 et 1940. Une modification des habitudes de vie et l'existence d'un dépistage organisé ou individuel pourraient expliquer ce phénomène. Parallèlement, le dépistage et une amélioration de la prise en charge thérapeutique pourraient expliquer la baisse de mortalité par cancer du sein observée pour les générations les plus récentes (*figure 2b*) [1].

**FIGURE 2a.** Courbes spécifiques de l'âge pour la cohorte 1930 – Échelle log-log – (d'après [1]).

**FIGURE 2b.** Risque cumulé 0-7 ans en pourcentage selon la cohorte de naissance (d'après [1]).



La *figure 3* représente les variations temporelles des taux d'incidence et de mortalité standardisés sur la population mondiale. Une augmentation de l'incidence est observée entre les années 1980 et 2005, avec des taux d'incidence standardisés (valeurs données pour 100 000 personnes-années) de 56,8 (pour l'année 1980), 82,0 (1995) et 101,5 (2005). Ces valeurs correspondent à une augmentation de l'incidence de 2,4 % par an (*figure 3*). Ces taux sont relativement élevés par rapport aux taux européens [3]. L'estimation française pour 2005 est voisine de l'estimation effectuée à partir du regroupement de neuf registres du programme SEER pour la population blanche américaine sur la période 1998-2002 (97,1 pour 100 000 personnes-années) [4]. La diminution des taux de mortalité par cancer du sein observée à partir de 1998 est estimée à - 1,3 % par an pour la période 2000-2005 [1]. L'augmentation temporelle de l'incidence associée à une diminution de la mortalité a été observée de façon assez générale en Europe entre les années 1995 et 2005 [3].

**FIGURE 3.** *Tendance chronologique en France (d'après [1]).*

En pratique, l'utilisation des taux standardisés permet d'effectuer des comparaisons géographiques et temporelles. Le calcul du risque cumulé par cohorte de naissance fournit une réelle appréciation du risque populationnel et de son évolution. Il s'agit bien évidemment d'un risque moyen. Les données fournies par les registres du cancer permettent d'estimer ce risque au niveau populationnel. L'analyse des données d'épidémiologie descriptive fournies par les registres du cancer permet l'étude des tendances évolutives du risque de cancer du sein au niveau populationnel.

## Risque relatif

Dans les études de cohorte, le risque relatif  $RR(t)$  peut être estimé, sous certaines conditions, à partir du taux relatif  $TR(t)$  défini comme le rapport des taux d'incidence spécifique à l'âge  $t$  des sujets exposés et non exposés au facteur de risque :

$$TR(t) = \frac{\ddot{e}_1(t)}{\ddot{e}_0(t)}$$

Dans le cas où le taux relatif ne dépend pas de l'âge, e.g.,  $TR(t) = TR$ , lorsque la maladie étudiée est suffisamment rare (ou la période d'observation suffisamment courte, de sorte que les valeurs des risques cumulés restent faibles), il est possible d'en déduire la relation entre les risques cumulés chez les individus exposés et non exposés [5] :

$$R_1(t) = 1 - e^{-E(t)} = 1 - \left\{ 1 - R_0(t) \right\}^{TR}$$

d'où l'on déduit :

$$TR \approx \frac{R_1(t)}{R_0(t)} = RR$$

Dans les études cas-témoins, le risque relatif est estimé à partir de l'odds-ratio. Par construction, les études cas-témoins ne permettent pas d'estimer le risque absolu mais uniquement le risque relatif.

## **Estimation du risque individuel : exemple du modèle de Gail**

Ce modèle a été développé à partir des données du Breast Cancer Detection Demonstration Project (BCDDP), une étude de cohorte incluant 243 221 femmes volontaires ayant bénéficié d'un dépistage annuel de cancer du sein (examen clinique et mammographie) sur la période 1973-1975. Une étude cas-témoins nichée dans cette cohorte a été réalisée pour estimer les risques relatifs associés aux facteurs de risque individuels. En combinant les taux d'incidence spécifiques par âge estimés à partir de la cohorte avec les risques relatifs estimés à partir de l'étude cas-témoins nichée dans celle-ci, le modèle fournit une estimation du risque de cancer du sein sur un intervalle d'âge donné prenant en compte les facteurs de risque individuels [6, 7].

## **Conclusion**

Les études de cohorte permettent de calculer l'estimation du risque (absolu) de développer un cancer sur un intervalle d'âge donné à partir des taux d'incidence spécifiques par âge. À partir des données d'incidence fournies par les registres de population, il est possible d'estimer ce risque pour une cohorte populationnelle donnée [1]. Bénichou et Gail [7] ont étudié les propriétés des estimations du risque individuel de développer un cancer sur un intervalle d'âge donné en combinant les estimations de l'incidence et celles du risque relatif.

## **Références bibliographiques**

- [1] Belot A, Grosclaude P, Bossard N et al. Cancer incidence and mortality in France over the period 1980-2005. *Rev Epidemiol Santé Publique* 2008;56:159-75.
- [2] Colonna M, Danzon A, Delafosse P et al. Cancer prevalence in France: time trend, situation in 2002 and extrapolation to 2012. *Eur J Cancer* 2008;44:115-22.
- [3] Karim-Kos HE, de Vries E, Soerjomataram I et al. Recent trends of cancer in Europe: a combined approach of incidence, survival and mortality for 17 cancer sites since the 1990's. *Eur J Cancer* 2008;55:1345-89.

- [4] Curado MP, Edwards B, Shin HR et al. *Cancer incidence in five continents (vol. IX)*. IARC scientific publications n° 160, Lyon, 1997.
- [5] Breslow NE, Day NE. *Statistical methods in cancer research (vol. I): the analysis of case-control studies*. IARC scientific publications n° 32, Lyon, 1980.
- [6] Gail MH, Bénichou J. *Estimating individualized risk of breast cancer*. In: *Cancer genetics for the clinician*, edited by Shaw. Kluwer Academic/Publishers Plenum, New-York, 1999.
- [7] Bénichou J, Gail MH. *Methods of inference for estimates of absolute risk derived from population-based case-control studies*. *Biometrics* 1995;51:182-94.