

Estimation de l'incidence départementale des cancers à partir de données médico-administratives*

Z Uhry¹, L Remontet², A Rogel¹, N Mitton^{3,4}, N Bossard², M Colonna^{3,4}

A Belot^{1,2}, P Grosclaude⁴, S Delacour-Billon⁴, J Gentil⁴, M Boussac¹, A Danzon¹,
M Altana⁵, F Frete⁵, A Weill⁵ (**Etude ALD**)

B Trombert⁶, F Olive⁶, F Gomez⁶, AM Schott⁶, S Polazzi⁶ (**Etude PMSI - ONCEPI**)

¹ Institut de veille sanitaire

² Service de Biostatistique, Hospices Civils de Lyon

³ Registre du cancer de l'Isère

⁴ Réseau des registres français du cancer Francim

⁵ Caisses d'Assurance Maladie (Caisse nationale des travailleurs salariés; Mutualité sociale agricole, Régime social des indépendants)

⁶ Départements d'information médicale (CHU Lyon, Grenoble et St-Etienne; Centre de lutte contre le cancer de Lyon)

* ALD: Affections de Longue Durée - données des caisses d'assurance maladie

* PMSI: Programme de Médicalisation des Systèmes d'Information - données hospitalières

Contexte

- Registres des cancers couvrent actuellement 20% de la population en France
- Estimation de l'incidence nationale réalisée en utilisant la mortalité comme corrélat de l'incidence
- En revanche, la mortalité ne peut être utilisée comme corrélat pour estimer l'incidence départementale
- Les bases médico-administratives (BMA) peuvent être utilisées comme corrélat de l'incidence pour estimer l'incidence départementale/régionale
- Nécessité d'évaluer la validité de telles estimations, pour chaque localisation cancéreuse

Objectifs

Pour les ALD et le PMSI séparément, en suivant une méthodologie commune :

- 1- Evaluer la validité des estimations de l'incidence départementale des cancers, à partir du rapport BMA/I observé dans les départements avec registres et des BMA de tous les départements**
⇒ Evaluation dans les départements avec registre, pour 24 localisations cancéreuses

- 2- Estimer l'incidence départementale des localisations cancéreuses, pour lesquelles la méthodologie aura été jugée valide**

Remarque : les deux études ALD et PMSI ont été menées séparément et pas au même moment
=> périodes d'étude ne sont pas identiques

Matériel

Etude ALD (France)

- Période 2000-2005
- Incidence Nombre de cas incidents (source registre, 8 à 11 depts selon cancer étudié)
- ALD Nombre de premières demandes d'ALD (source assurance maladie*, tous depts)
- Format Données analysées par localisation cancéreuse, sexe, département et âge quinquennal
- Notation ALD/I : rapport entre le nombre de première demande d'ALD et le nombre de cas incidents

Etude PMSI (France)

- Période 2004-2007
- Incidence Nombre de cas incidents (source registre, 8 à 11 depts selon cancer étudié)
- PMSI Nombre de patients hospitalisés sur la période avec un diagnostic principal de cancer (source base nationale hospitalière, tous depts)
Exclusion des départements avec un taux de chaînage patient < 90% et redressement sinon
- Format Données analysées par localisation cancéreuse, sexe, département et âge quinquennal
- Notation PMSI/I : rapport entre le nombre de patients hospitalisés et le nombre de cas incidents

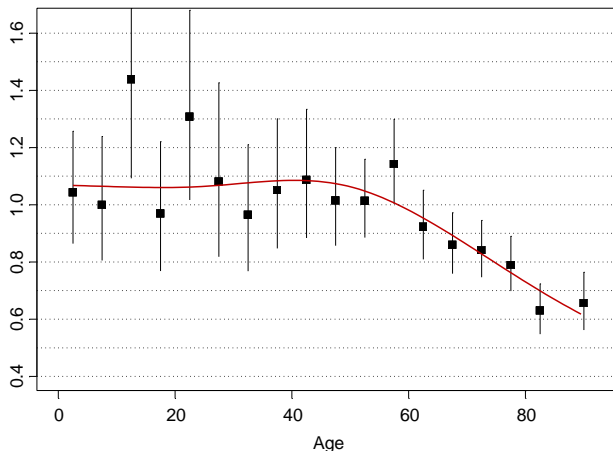
*: trois principaux régime (régime général, régime agricole et régime des indépendants)

■ Méthode d'estimation

- **Modélisation du rapport ALD/I**, en fonction de l'âge et du département (j), **depts avec registre**, par un modèle de Poisson à effet mixtes:

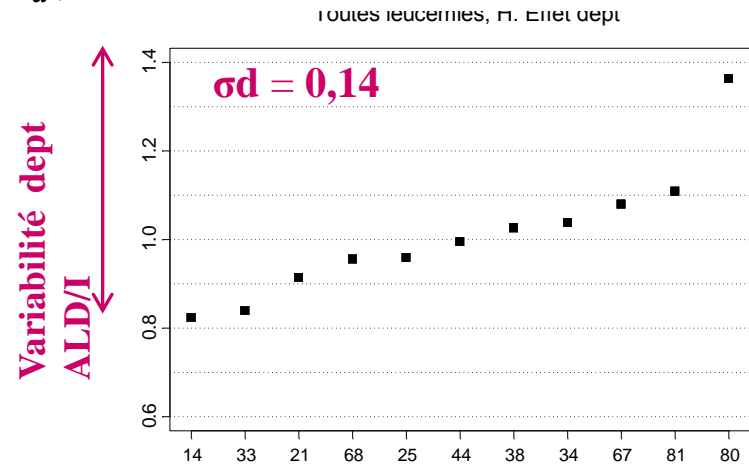
$$(1): (ALD/I)_{age, j} = f(age).exp(b_j), \quad \text{où } b_j \sim N(0, \sigma_d^2)$$

• $f(age)$: rapport ALD/I selon l'âge



• $exp(b_j)$: effet dept (ie écart / moyenne des depts)

• σ_d : variabilité départementale du rapport ALD/I



- **Estimations du nombre de cas incidents par département (96 départements):**

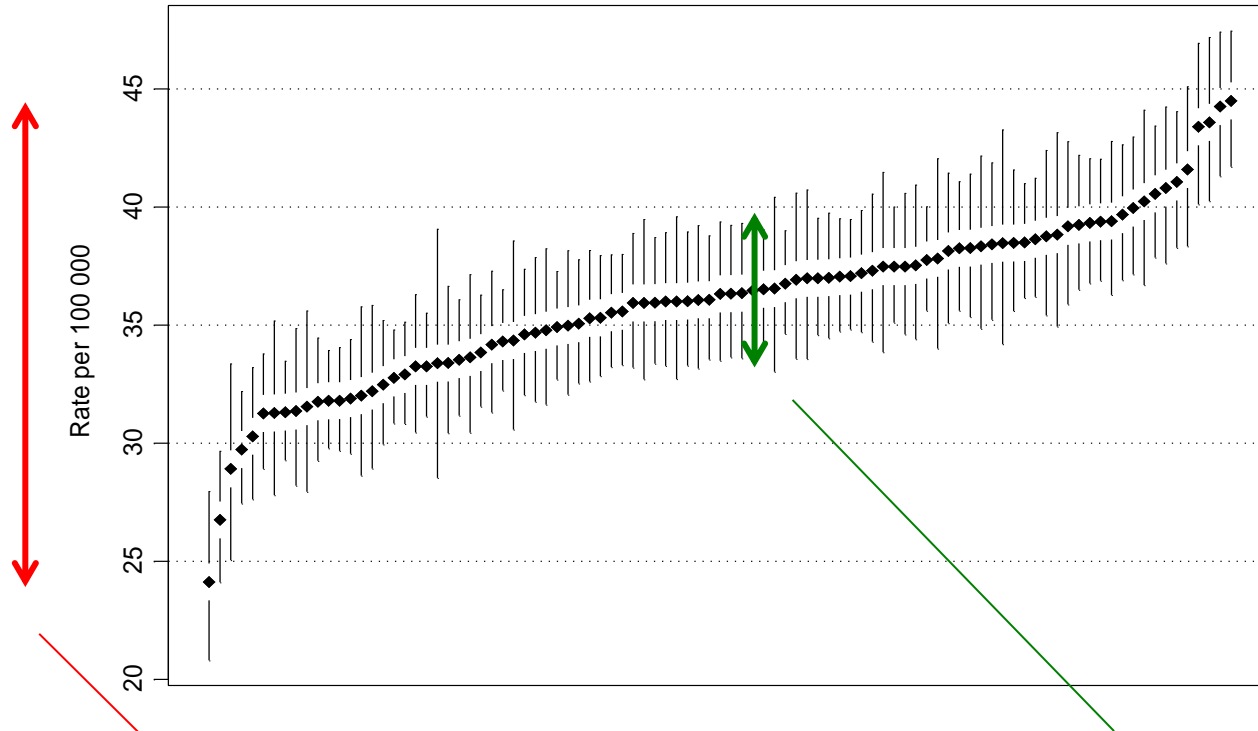
$$\hat{I}_{age,dept} = ALD_{age,dept} / f(age) \quad ; \sigma_d \text{ intervient dans calcul IC}$$

- **Hypothèse:** rapport ALD/I, à âge égal, identique dans tous les depts (σ_d petit)

■ Validation de l'utilisation du rapport ALD/I pour estimer l'incidence départementale des cancers

Evaluation dans les départements avec registres, à partir de :

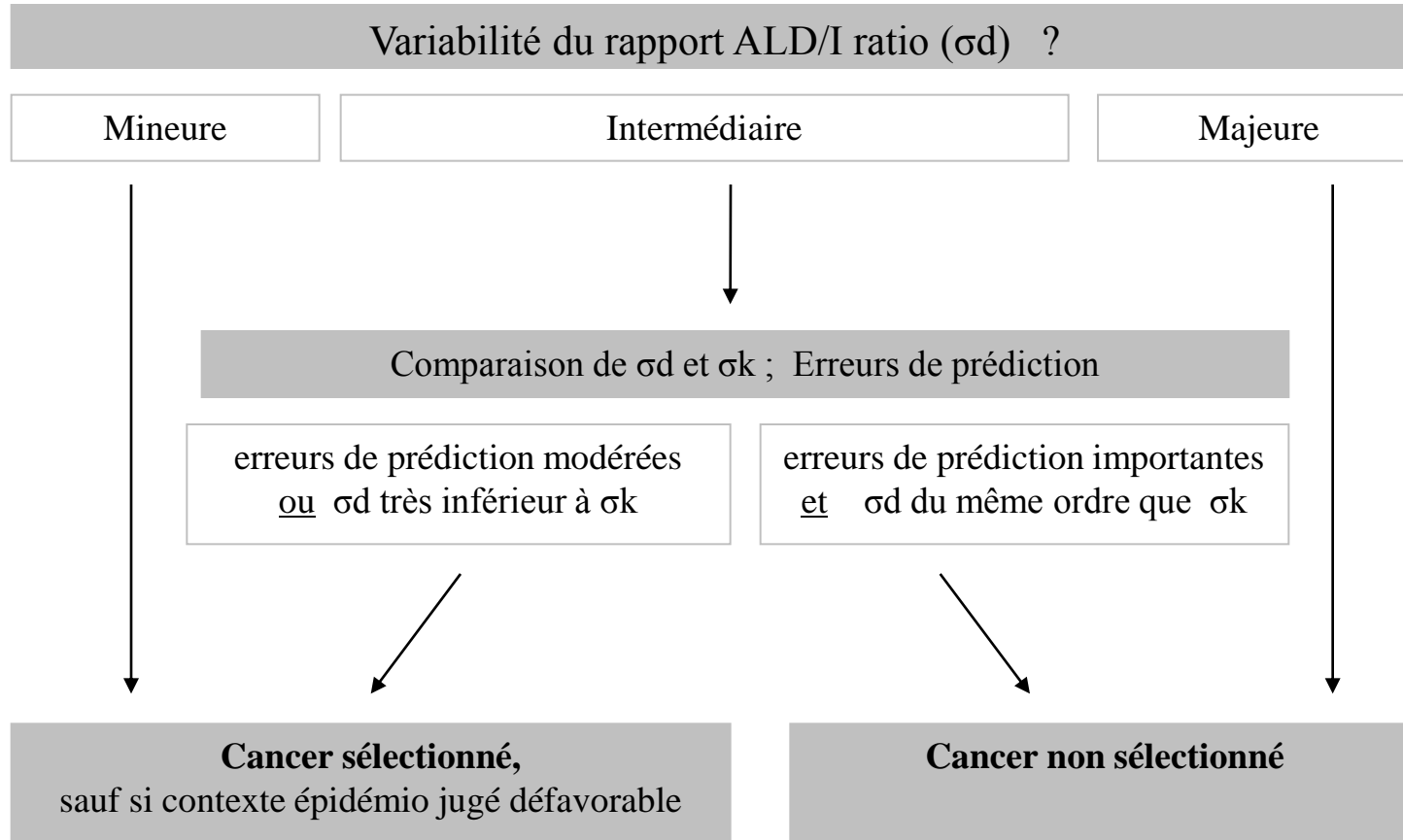
- **Mesure de la variabilité départementale du rapport ALD/I** : σ_d issu modèle (1)
- **Mesure des erreurs de prédiction** (comparaison de l'incidence observée et prédite) : **validation croisée**
 - Pour chaque dept, nombre de cas incidents est prédit en utilisant le rapport ALD/I obtenu en excluant ce dept
 - Comparaison des nombres observés et prédits (mesure du type $\text{Chi}^2 : (O-P)^2/P$) et erreur relative ($ER=O/P - 1$)
- **Mesure de la variabilité départementale de l'incidence observée**
 - Mesurée par un modèle du même type que (1), où la variable expliquée est cette fois le taux d'incidence $\Rightarrow \sigma_k$
 - Comparaison de σ_d et de σ_k : qualité des estimations comparée à l'ampleur des variations géo d'incidence
- **Contexte épidémiologique**
- **Et, pour le PMSI, variabilité départementale du % de double tumeurs (par localisation)**



**Ampleur des variations de l'incidence,
estimée par σ_k**

**Qualité (ie précision) de l'estimation, liée
à σ_d**

■ Principe de sélection des localisations cancéreuses



⇒ **Discussion** au sein du groupe de travail pour chaque localisation

- **Estimations de l'incidence des cancers dans les 96 départements**
 - Pour les localisations considérées comme valides à l'issue de l'étape de validation
 - Selon méthode décrite plus haut
 - Estimations des nombres de cas, taux standardisés et rapports standardisé d'incidence (SIR)
 - Fournies avec intervalle de prédiction (IP) à 95%, qui prennent en compte σ_d ** important **

Dept	LBP - Hommes $\sigma d = 0,03$; $\sigma k = 0,22$			Leucémies - Hommes $\sigma d = 0,14$; $\sigma k = 0,14$		
	Obs	Pred	ER (%)	Obs	Pred	ER (%)
Calvados	790	756	-4	274	210	* -24
Côte d'Or ^S	-	-	-	252	222	* -12
Doubs	489	525	7	239	225	-6
Gironde ^S	-	-	-	446	349	* -22
Hérault	854	796	* -7	420	440	5
Isère	942	909	-3	493	511	4
Loire-Atl.	1404	1501	* 7	665	670	1
Bas-Rhin	1124	1091	-3	435	488	* 12
Haut-Rhin	657	706	8	351	328	-7
Somme	807	822	2	181	273	* 51
Tarn	283	265	-6	167	190	13
Total	7 350	7 372	0	3 923	3 905	0

ER: erreur relative (O/P-1) ; * p< 0.05

LBP: Lèvre-Bouche-Pharynx ; ^S: Registres spécialisés en hématologie

Résultats (2) – Bilan de l'étape de validation, ALD et PMSI

⇒ 8 localisations cancéreuses retenues pour ALD et 9 pour le PMSI, dont 6 en commun

Localisation cancéreuse	ALD (2000-05)	PMSI (2004-07)
Lèvre-bouche-pharynx (LBP)	H / F	
Esophage	H *	
Estomac	H / F	H / F
Côlon-rectum	H / F	H / F
Foie		
Pancréas		
Larynx		H *
Poumon	H / F	H / F
Mélanome		
Sein	F	F
Col utérus		F
Corps utérus		F
Ovaire	F	F
Prostate		
Testicule	H	H
Vessie		
Rein		
Système nerveux central		
Thyroïde		
Lymphome malin non		
Maladie de Hodgkin		
Myélome multiple		
Toutes Leucémies		

* Effectifs trop faibles pour analyse chez la femme

Discussion

- L'évaluation est un préalable indispensable à l'utilisation des BMA à des fins de surveillance
- Cette évaluation est limitée par le nombre de départements avec registre
- Elle ne peut pas reposer sur un critère purement statistique: une discussion est nécessaire pour chaque cancer
- L'expertise épidémiologique et des professionnels (DIM, assurance maladie) est fondamentale
- Dans un premier temps, les ALD et le PMSI ont été évaluées séparément : une confrontation est prévue sur une même période
- Seul un tiers des localisations étudiées ont été retenues pour chaque BMA: il est difficile d'obtenir des estimations départementales valides à partir des BMA
- Des progrès sont toutefois possibles...

Perspectives

- Le croisement individuel des données ALD et PMSI est maintenant possible et sera également évalué
 - Plusieurs sources (PMSI, ALD, croisement ALD-PMSI et/ou mortalité) sont disponibles; elles peuvent être utilisées seules ou conjointement
- ⇒ Multiples possibilités: il faudra déterminer la meilleure option, pour chaque localisation cancéreuse